

CENTRAL SWATHS

(A GENERALIZATION OF THE CENTRAL PATH)

JAMES RENEGAR

ABSTRACT. We develop a natural generalization to the notion of the central path – a notion that lies at the heart of interior-point methods for convex optimization. The generalization is accomplished via the “derivative cones” of a “hyperbolicity cone,” the derivatives being direct and mathematically-appealing relaxations of the underlying (hyperbolic) conic constraint, be it the non-negative orthant, the cone of positive semidefinite matrices, or other.

We prove that a dynamics inherent to the derivative cones generates paths always leading to optimality, the central path arising from a special case in which the derivative cones are quadratic. Derivative cones of higher degree better fit the underlying conic constraint, raising the prospect that the paths they generate lead to optimality quicker than the central path.

1. Introduction

Let \mathcal{E} denote a finite-dimensional Euclidean space and let $p : \mathcal{E} \rightarrow \mathbb{R}$ be a hyperbolic polynomial, that is, a homogeneous polynomial for which there is a designated direction vector e satisfying $p(e) > 0$ and having the property that for all $x \in \mathcal{E}$, the univariate polynomial $t \mapsto p(x + te)$ has only real roots. Thus, p is “hyperbolic in direction e .”

Let Λ_{++} denote the hyperbolicity cone – the connected component of $\{x : p(x) > 0\}$ containing e . Let Λ_+ be the closure.

A simple example is $p(x) = \prod_j x_j$ and $e = (1, \dots, 1)$, in which case Λ_{++} is the strictly positive orthant and Λ_+ is the non-negative orthant. Perhaps the most fundamental example, however, is $p(X) = \det(X)$, where X ranges over $n \times n$ symmetric matrices and $e = I$, the identity matrix. Here, Λ_{++} is the cone of (strictly-)positive definite (pd) matrices and Λ_+ is the positive semidefinite (psd) cone.

Gårding [6] showed for each hyperbolic polynomial p that every $\hat{e} \in \Lambda_{++}$ is a hyperbolicity direction, i.e., for every x , all of the roots of $t \mapsto p(x + t\hat{e})$ are real. One of several remarkable corollaries Gårding established is that Λ_{++} is convex. (Of course Λ_+ is thus convex, too.) (See §2 of [13] for simplified proofs.)

The combination of convexity and rich algebraic structure make hyperbolicity cones promising objects for study in the context of optimization, as was first made evident by Güler [7], who developed a rich theory of interior-point methods for

1991 *Mathematics Subject Classification.* 90C05, 90C22, 90C25, 52A41, 52B15.

Key words and phrases. hyperbolicity cone, hyperbolic polynomial, hyperbolic programming, central path, conic programming, convex optimization.

Research supported in part by NSF Grant #CCF-0430672.

Thanks to Chek Beng Chua and Yuriy Zinchenko for many helpful conversations. Deep gratitude also goes to the referees, whose extensive comments led to the paper being significantly restructured, and led to the (motivational) exposition being considerably expanded especially in the proofs.

hyperbolic programs, that is, for problems of the form

$$\begin{aligned} \min \quad & c^*x \\ \text{s.t.} \quad & Ax = b \\ & x \in \Lambda_+ \end{aligned}$$

– linear programming, second-order programming and semidefinite programming being particular cases. Key to Güler’s development is that the function $x \mapsto -\ln p(x)$ is a self-concordant barrier for Λ_+ ; thus the general theory of Nesterov and Nemirovski [12] applies.

A primary purpose of the present paper is to use the viewpoints provided by hyperbolic programming to develop a natural generalization to the notion of the central path¹ (a notion that lies at the heart of interior-point method theory). This is accomplished via derivative cones, which are direct relaxations of the underlying convex conic constraint (be it the non-negative orthant, the cone of positive semidefinite matrices, ...).

However, perhaps more important than the “natural generalization to the notion of the central path” is our “use (of) the viewpoints provided by hyperbolic programming” in developing the generalization. Indeed, it is our conviction that even if results about hyperbolic programming never find application more general than linear programming and semidefinite programming, the setting of hyperbolic programming is favorable for engendering intriguing algorithmic ideas that otherwise would have been unrealized (or at least considerably delayed).

Familiarity with the central path is not required to readily understand our results. (The central path simply provides an initial anchor with which many readers *are* familiar.)

The literature focusing on hyperbolic polynomials is relatively small but its growth is accelerating and its quality in general is distinctly impressive. Although the nature of our results is such that during the development we have occasion to cite only a few works, we take the opportunity before beginning the development to draw the reader’s attention to the bibliography, which includes a variety of notable papers appearing in recent years. In particular, an appreciation of the breadth and quality of research ideas surrounding hyperbolic polynomials can be fostered by browsing [1], [2], [8], [9], and [10].

2. Overview of Results

Let ϕ be a univariate polynomial all of whose coefficients are real. Between any two real roots of ϕ there lies, of course, a root of ϕ' . Consequently, because ϕ' is of degree one less than the degree of ϕ , a simple counting argument shows that if all of the roots of ϕ are real, then so are all of the roots of ϕ' .

In particular, if $\phi(t) := p(x + te)$ where p is a polynomial hyperbolic in direction e (and where x is an arbitrary point), then all the roots of $t \mapsto \phi'(t) = \frac{d}{dt}p(x + te) = Dp(x + te)[e]$ are real, where $Dp(x + te)$ denotes the differential of p at $x + te$. Hence, the polynomial $p'_e(x) := Dp(x)[e]$ is, like p , hyperbolic in direction e .

For example, if $p(x) = \prod_j x_j$ and all coordinates of e are nonzero, then $p'_e(x) = \sum_i e_i \prod_{j \neq i} x_j$ is hyperbolic in direction e .

We refer to p'_e as the “derivative polynomial (in direction e),” and denote its hyperbolicity cone by Λ'_{++e} . The fact that for every x , the roots of $t \mapsto p'_e(x + te)$ lie between the roots of $t \mapsto p(x + te)$ is readily seen to imply $\Lambda_{++} \subseteq \Lambda'_{++e}$ – in words, Λ'_{++e} is a *relaxation* of Λ_{++} (see §4 of [13] for a full discussion).

Of course one can in turn take the derivative in direction e of the hyperbolic polynomial p'_e , thereby obtaining yet another polynomial – $(p'_e)'_e(x) = D^2p(x)[e, e]$

¹Central Path := $\{x(\eta) : \eta > 0\}$ where $x(\eta)$ solves $\min_x \eta c^*x - \ln p(x)$, s.t. $Ax = b$, $x \in \Lambda_{++}$.

– hyperbolic in direction e . Letting n denote the degree of p , repeated differentiation in direction e results in a sequence of hyperbolic polynomials

$$p_e^{(1)} = p'_e, p_e^{(2)}, \dots, p_e^{(n-1)},$$

where $\deg(p_e^{(i)}) = n - i$. (For convenience, let $p_e^{(0)} := p$.) The associated hyperbolicity cones $\Lambda_{+,e}^{(i)}$ – and their closures $\Lambda_{+,e}^{(i)}$ – form a nested sequence of relaxations of the original cone:

$$\Lambda_+ = \Lambda_{+,e}^{(0)} \subseteq \Lambda_{+,e}^{(1)} \subseteq \Lambda_{+,e}^{(2)} \subseteq \dots \subseteq \Lambda_{+,e}^{(n-1)}.$$

The final relaxation, $\Lambda_{+,e}^{(n-1)}$, is a halfspace, because $p_e^{(n-1)}$ is linear.

The cones become tamer as additional derivatives are taken. The halfspace $\Lambda_{+,e}^{(n-1)}$ is as tame as a cone can be, but extremely tame also is the second-order cone $\Lambda_{+,e}^{(n-2)}$ – no cone with curvature could be nicer. As one moves along the nesting towards the original cone Λ_+ , the boundaries gain more and more corners. For example, when $p(x) = \prod_j x_j$ and all coordinates of e are positive, the boundary $\partial\Lambda_{+,e}^{(i)}$ contains all of the non-negative orthant's faces of dimension less than $n - i$ (hence *lots* of corners when i is small and n large). On the other hand, everywhere else, $\partial\Lambda_{+,e}^{(i)}$ has nice curvature properties (no corners), as is reflected in the following motivational theorem pertaining to every hyperbolicity cone whose closure is regular (i.e., has nonempty interior and contains no subspace other than the origin).

Theorem 1. *Assume Λ_+ is regular and $0 \leq i \leq n - 2$.*

- (A) *The intersection $\Lambda_+ \cap \partial\Lambda_{+,e}^{(i)}$ is independent of $e \in \Lambda_{++}$ (thus, a face of Λ_+ which is a boundary face of $\Lambda_{+,e}^{(i)}$ for some $e \in \Lambda_{++}$ is a boundary face for all $e \in \Lambda_{++}$).*
- (B) *If $e \in \Lambda_{++}$ then any boundary face of $\Lambda_{+,e}^{(i)}$ either is a face of Λ_+ or is a single ray contained in $\Lambda_{+,e}^{(i+1)} \setminus \Lambda_{+,e}^{(i)}$.*

We show in §4 that the theorem is a consequence of results from [13]. (In order to make the present section inviting to a broad audience, nearly all proofs are delayed.)

Before moving to discussion of hyperbolic programs, we record a characterization of the derivative cones that is useful both conceptually and in proofs:

$$\Lambda_{+,e}^{(i)} = \{x : p_e^{(j)}(x) \geq 0 \text{ for all } j = i, \dots, n - 1\}. \quad (2.1)$$

(This is immediate from Proposition 18 and Theorem 20 of [13].)



Consider a hyperbolic program

$$\left. \begin{array}{ll} \min & c^*x \\ \text{s.t.} & Ax = b \\ & x \in \Lambda_+ \end{array} \right\} \text{HP}$$

and its derivative relaxations in direction e ,

$$\left. \begin{array}{ll} \min & c^*x \\ \text{s.t.} & Ax = b \\ & x \in \Lambda_{+,e}^{(i)} \end{array} \right\} \text{HP}_e^{(i)} \quad (i = 1, \dots, n - 1).$$

(Strictly speaking, “min” should be replaced with “inf,” but we focus on instances where a minimizer exists.) The optimal values for the derivative relaxations $\text{HP}_e^{(i)}$ form a decreasing sequence in i , due to the nesting of the derivative cones.

Let Feas (resp., $\text{Feas}_e^{(i)}$) denote the feasible region of HP (resp., $\text{HP}_e^{(i)}$) – the set of points satisfying the constraints. Let Opt ($\text{Opt}_e^{(i)}$) denote the set of optimal points – a.k.a. optimal solutions – and let val denote the optimal value of HP.

We assume $b \neq 0$ (thus, the origin is infeasible, and the feasible sets are not cones), A is surjective (i.e., onto), and c^* is not in the image of A^* (otherwise every feasible point would be optimal).

We assume Λ_+ is a regular cone. Then, for $1 \leq i \leq n-2$, $\Lambda_{+,e}^{(i)}$ also is regular ([13], Proposition 13).

From these assumptions and Theorem 1(B) immediately follows a fact that will play a critical role:

$$\left. \begin{array}{l} \text{If } 1 \leq i \leq n-2, \text{ then either} \\ \text{Opt}_e^{(i)} = \emptyset, \\ \text{Opt}_e^{(i)} = \text{Opt}, \text{ or} \\ \text{Opt}_e^{(i)} \text{ consists of a single point} \\ \text{and the point is contained in } \text{relint}(\text{Feas}_e^{(i+1)}) \setminus \text{Feas}_e^{(i-1)}, \end{array} \right\} \quad (2.2)$$

where “relint” denotes relative interior².

Thus, for each $1 \leq i \leq n-2$, the cone Λ_{++} is naturally partitioned into three regions, one consisting of derivative directions e for which $\text{Opt}_e^{(i)} = \text{Opt}$, a second consisting of directions for which $\text{Opt}_e^{(i)} = \emptyset$, and the third consisting of directions for which $\text{Opt}_e^{(i)}$ consists of a single point lying outside the feasible region for the original optimization problem HP. We associate names with this partitioning of Λ_{++} , but before doing so, we introduce a restriction.

We shall only be concerned with derivative directions e satisfying $Ae = b$ (indeed, key arguments rely heavily on $A(e-x) = 0$ for $x \in \text{Feas}_e^{(i)}$). Thus, the derivative directions we consider satisfy $e \in \text{Feas}_e^{(i)}$ for all i – in particular, $\text{HP}_e^{(i)}$ is “strictly” feasible, as $e \in \Lambda_{++} \subseteq \Lambda_{+,e}^{(i)}$.

We distinguish two sets of derivative directions for $0 \leq i \leq n-1$:

The i^{th} *central swath* is the set

$$\text{Swath}(i) := \{e \in \Lambda_{++} : Ae = b \text{ and } \text{Opt}_e^{(i)} \neq \emptyset\},$$

and the set of *core* derivative directions is defined by

$$\text{Core}(i) := \{e \in \text{Swath}(i) : \text{Opt}_e^{(i)} = \text{Opt}\}.$$

When $1 \leq i \leq n-2$ and $e \in \text{Swath}(i) \setminus \text{Core}(i)$, we use $x_e^{(i)}$ to denote the unique point in $\text{Opt}_e^{(i)}$ (unique by (2.2)).

For reference, we note that from (2.2),

$$x_e^{(i)} \in \Lambda_{+,e}^{(i+1)} \cap \partial\Lambda_{+,e}^{(i)}. \quad (2.3)$$

Trivially, if $\text{Opt} \neq \emptyset$ (resp., $= \emptyset$), then $\text{Swath}(0) = \text{Core}(0) = \text{relint}(\text{Feas})$ (resp., $= \emptyset$), that is, the zeroth swath coincides precisely with the relative interior of HP’s feasible region (resp., is the empty set). More interestingly, the swaths and cores are nested:

$$\begin{aligned} \text{Swath}(0) &\supseteq \text{Swath}(1) \supseteq \cdots \supseteq \text{Swath}(n-1) \\ \text{Core}(0) &\supseteq \text{Core}(1) \supseteq \cdots \supseteq \text{Core}(n-1) \end{aligned}$$

For the cores, the nesting is an easy consequence of the reverse nesting

$$\text{Feas}_e^{(0)} \subseteq \text{Feas}_e^{(1)} \subseteq \cdots \subseteq \text{Feas}_e^{(n-1)}.$$

These reverse nesting also provide the crux in proving the nesting of the swaths, a proof we defer to §4.

²The “relative interior” of a convex set $S \subseteq \mathcal{E}$ is the interior of S when considered as a subset in the smallest affine space containing S (where the affine space inherits the topology of the Euclidean space \mathcal{E}).

A consequence of the nesting of swaths is that if any swath is nonempty, then $\text{Swath}(0) \neq \emptyset$ – equivalently, $\text{Opt} \neq \emptyset$.

Whereas a path is narrow (one-dimensional), swaths can be broad, just as $\text{Swath}(i)$ typically fills much of the feasible region for HP when i is small. But why do we use the terminology “*central swaths*” rather than simply “swaths”? The following elementary theorem (proven in §4) gives our first reason.

Theorem 2. $\text{Swath}(n-1) = \text{Central Path}$.

The central path is fundamental in the literature on interior-point methods. The path leads to optimality. Most of the algorithms follow the path, either explicitly or implicitly. A foremost goal of the present paper is to show that not only does the central path lead to optimality, but *all* central swaths lead, in a natural manner, to optimality. We show, in particular, that through each point $e \in \text{Swath}(i)$ (for $1 \leq i \leq n-2$), there is naturally defined a trajectory which leads from e to optimality; moreover, the trajectory remains within $\text{Swath}(i)$ until optimality is reached. An intriguing possibility is that for small values of i , the trajectory might lead to optimality “more quickly” than the central path. (Motivation for this possibility will become clearer as the reader proceeds.)

For $1 \leq i \leq n-2$, consider the idealized setting in which for derivative directions $e \in \text{Swath}(i)$, an exact optimal solution for $\text{HP}_e^{(i)}$ can be computed. If the optimal solution lies in Λ_+ , then clearly it lies in Opt , the set of optimal solutions for the original optimization problem HP. In this case our goal of solving HP has been accomplished. On the other hand, if the optimal solution does not lie in Λ_+ , then $e \in \text{Swath}(i) \setminus \text{Core}(i)$, and the optimal solution is the unique point $x_e^{(i)}$ in $\text{Opt}_e^{(i)}$. In this case how can we move towards solving HP? How can we construct a trajectory $t \mapsto e(t)$ for which $e(0) = e$ and such that either the trajectory converges to Opt or the path $t \mapsto x_{e(t)}^{(i)}$ converges to Opt (or both)?

An apparently easier task would be to create a trajectory $t \mapsto e(t)$ for which $t \mapsto c^*e(t)$ is monotonically decreasing. Indeed, we could define the trajectory implicitly according to the differential equation $\frac{d}{dt}e(t) = x_{e(t)}^{(i)} - e(t)$ ($e(0) \in \text{Swath}(i) \setminus \text{Core}(i)$) that is, move from $e(t)$ infinitesimally towards the optimal solution $x_{e(t)}^{(i)}$. Assuming this does result in a well-defined trajectory, then clearly, $t \mapsto c^*e(t)$ is decreasing. However, there are no clear reasons suggesting that the trajectory $t \mapsto e(t)$ converges to Opt . It is conceivable, for example, that the trajectory reaches the boundary $\partial\Lambda_+$ in finite time, converging to a point having better objective value than $e(0)$, but not to a point in Opt . Alternatively, in finite time the trajectory might reach $\text{Core}(i)$. It seems plausible that the path $t \mapsto x_{e(t)}^{(i)}$ then would have limit in Opt . But how would one prove it? How does one even rule out the possibility that in finite time, the path $t \mapsto x_{e(t)}^{(i)}$ goes to infinity while the trajectory $t \mapsto e(t)$ remains bounded but with no limit points in Opt ?

Resolving these issues, and similar ones, is our primary focus. We show that the differential equation $\frac{d}{dt}e(t) = x_{e(t)}^{(i)} - e(t)$, $e(0) \in \text{Swath}(i) \setminus \text{Core}(i)$, does result in well-defined trajectories in $\text{Swath}(i) \setminus \text{Core}(i)$, and we show that either the trajectory $t \mapsto e(t)$ or the path $t \mapsto x_{e(t)}^{(i)}$ does converge to Opt . We show many other things as well, but to accurately explain, first we must formalize.

In place of $\frac{d}{dt}e(t) = x_{e(t)}^{(i)} - e(t)$ we often write $\dot{e}(t) = x_{e(t)}^{(i)} - e(t)$. That this dynamics results in well-defined trajectories is immediate from the following theorem, whose (relatively routine) proof is in §5.

Theorem 3. *Assume $1 \leq i \leq n-2$. The set $\text{Swath}(i) \setminus \text{Core}(i)$ is open in the relative topology of $\text{relint}(\text{Feas})$. Moreover, the map $e \mapsto x_e^{(i)}$ is analytic on $\text{Swath}(i) \setminus \text{Core}(i)$.*

As an aside, we note that every $e \in \text{Swath}(n-2)$ has a unique optimal solution, simply due to the strict curvature of second-order cones. Thus, we can naturally extend the definition of $x_e^{(n-2)}$ to include all derivative directions in $\text{Swath}(n-2)$, not just the ones in $\text{Swath}(n-2) \setminus \text{Core}(n-2)$. It happens that for the case $i = n-2$, virtually all of our results remain valid when “ $\text{Swath}(n-2)$ ” is substituted for “ $\text{Swath}(n-2) \setminus \text{Core}(n-2)$ ” (as we discuss while proving our theorems (§§3-10)). With regards to Theorem 3 in particular, the extended map $e \mapsto x_e^{(n-2)}$ is analytic on all of $\text{Swath}(n-2)$ (which is open in the relative topology of $\text{relint}(\text{Feas})$).

Theorem 3 implies that when initiated at $e(0) \in \text{Swath}(i) \setminus \text{Core}(i)$, the dynamics $\dot{e}(t) = x_{e(t)}^{(i)} - e(t)$ results in a well-defined trajectory. The trajectory remains in $\text{Swath}(i) \setminus \text{Core}(i)$ for all time (i.e., is defined for all $0 \leq t < \infty$), or is defined only up to some finite time due either to reaching the boundary of $\text{Swath}(i) \setminus \text{Core}(i)$ or escaping to infinity. Let $T(e(0))$ denote the time at which the trajectory becomes undefined (possibly $T(e(0)) = \infty$). We refer to $t \mapsto e(t)$ ($0 \leq t < T(e(0))$) as a “maximal trajectory.” For brevity, we often instead write “ $0 \leq t < T$ ” with the implicit understanding that the time of termination, T , depends on $e(0)$. (Some of our results distinguish between $T < \infty$ and $T = \infty$, but never is distinction made between different finite termination times.)

Here is the formal statement of results partially described earlier:

Main Theorem (Part I). *Assume $1 \leq i \leq n-2$ and let $t \mapsto e(t)$ ($0 \leq t < T$) be a maximal trajectory generated by the dynamics $\dot{e}(t) = x_{e(t)}^{(i)} - e(t)$ beginning at $e(0) \in \text{Swath}(i) \setminus \text{Core}(i)$.*

- (A) *The trajectory $t \mapsto e(t)$ is bounded, and $t \mapsto c^*x_{e(t)}^{(i)}$ is strictly increasing, with val (the optimal value of HP) as the limit.*
- (B) *If $T = \infty$ then every limit point of the trajectory $t \mapsto e(t)$ lies in Opt .*
- (C) *If $T < \infty$ then the trajectory $t \mapsto e(t)$ has a unique limit point \bar{e} and $\bar{e} \in \text{Core}(i)$; moreover, the path $t \mapsto x_{e(t)}^{(i)}$ is bounded and each of its limit points lies in Opt .*

The Main Theorem (Part I) is proven in §8.

An immediate consequence of the theorem is that $T = \infty$ whenever $\text{Core}(i) = \emptyset$, as is the case whenever $\text{Opt} \cap \Lambda_{++e}^{(i)} \neq \emptyset$ for some (and hence, by Theorem 1(A), for all) $e \in \Lambda_{++}$.

Perhaps the reader wonders as to the inspiration for the idea that the trajectories $t \mapsto e(t)$ arising from the dynamics $\dot{e}(t) = x_{e(t)}^{(i)} - e(t)$ lead to optimality, either in the limits of the trajectories themselves or in the limits of the paths $t \mapsto x_{e(t)}^{(i)}$. The following theorem (whose proof is in §6) serves to clarify the inspiration, as well as to further illuminate our choice of the terminology “central swaths” (as opposed to simply “swaths”).

Theorem 5. *Assume $t \mapsto e(t)$ ($0 \leq t < T$) is a maximal trajectory arising from the dynamics $\dot{e}(t) = x_{e(t)}^{(n-2)} - e(t)$, starting at $e(0) \in \text{Swath}(n-2) \setminus \text{Core}(n-2)$.*

If $e(0) \in \text{Central Path}$ then $\{e(t) : 0 \leq t < T\} \subseteq \text{Central Path}$.

We remark that when $i = n-2$, the termination time T always is ∞ , even when $\text{Core}(n-2) \neq \emptyset$ (see §6).

From the two theorems, we see that the central path is but one trajectory in a rich spectrum of paths. Moreover, the central path is at the far end of the spectrum, where the cone Λ_+ is relaxed to second-order cones $\Lambda_{+,e}^{(n-2)}$. Second-order cones have nice curvature properties but are far cruder approximations to Λ_+ than are cones $\Lambda_{+,e}^{(i)}$ for small i . This raises the interesting prospect that algorithms more efficient than interior-point methods can be devised by relying on a smaller value of i , or on a range of values of i in addition to $i = n - 2$.

Some exploration in this vein has been made by Zinchenko ([14],[16]), who showed for linear programs satisfying standard non-degeneracy conditions that if i is chosen appropriately and the initial derivative direction $e(0)$ is within a certain region, then a particular algorithm based on discretizing the flow $\dot{e}(t) = x_{e(t)}^{(i)} - e(t)$ converges R-quadratically to an optimal solution.

Before moving to the next result, we acknowledge that maybe the Main Theorem can be strengthened without restricting its general setting. For example, in Part I(B) there is no statement that limit points of the path $t \mapsto x_{e(t)}^{(i)}$ are optimal solutions for HP – there is not even a statement that the path is bounded. This omission seems odd given that the trajectory $t \mapsto e(t)$ is following the path $t \mapsto x_{e(t)}^{(i)}$ (according to the dynamics $\dot{e}(t) = x_{e(t)}^{(i)} - e(t)$) and given that the theorem states limit points of the trajectory are optimal for HP. Intuitively, it seems the path would converge to optimality and do so even more quickly than the trajectory. The intuition is correct for a wide variety of “non-degenerate” problems (indeed, quicker convergence of the path than the trajectory underlies Zinchenko’s speedup), but we have been unable to find a proof – or counterexample – in the general setting of the theorem.³

In a similar vein, maybe it is true when $T = \infty$ that the trajectory $t \mapsto e(t)$ has a *unique* limit point. If, like the central path, the trajectory was a semialgebraic set then the limit point indeed would be unique (simply because every semialgebraic path that has a limit point has exactly one limit point). However, we doubt that the trajectories are necessarily semialgebraic in general, and we see no other approach to proving uniqueness. The theorem leaves open for the general setting the possibility that when $T = \infty$ (resp., $T < \infty$), some trajectories $t \mapsto e(t)$ (resp., some paths $t \mapsto x_{e(t)}^{(i)}$) have non-trivial limit cycles – yet we have no examples of such behavior.



Of course the dynamics of moving e towards an optimal solution x can also be done for $e \in \text{Core}(i)$, in which case e would be moving towards $x \in \text{Opt}_e^{(i)} = \text{Opt}$. As a matter of formalism, it would be nice to know that such movement would result in a new derivative direction for which x is still optimal, that is, a new derivative direction that also is in $\text{Core}(i)$. The following theorem (proven in §10) establishes a bit more.

Theorem 6. *Assume $e \in \text{Core}(i)$ and let \mathcal{A} be the minimal affine space containing both e and Opt . Then*

$$\mathcal{A} \cap \Lambda_{++} \subseteq \text{Core}(i) .$$

In the following conjecture, the empty set is taken, by default, to be convex.

Conjecture. *$\text{Core}(i)$ is convex.*

³To gain a sense of the difficulties (and how first impressions can mislead), consider that for any value $0 < T \leq \infty$, it is straightforward to define dynamics on \mathbb{R}^2 that generates a pair of paths $a(t)$, $b(t)$ for which $\dot{b}(t) = a(t) - b(t)$ and as $t \rightarrow T$, $a(t)$ spirals outward to infinity whereas $b(t)$ spirals inward to a point. (Thus, although $a(t)$ is “leading” $b(t)$, the paths end (infinitely) far apart.)



Much work remains in order to transform the ideas captured in the Main Theorem (Part I) into general and efficient algorithms. For example, devising and analyzing efficient methods for computing $x_e^{(i)}$ given e is, in the general case, a challenging research problem. However, computing $x_e^{(n-2)}$ (that is, the case $i = n - 2$) amounts simply to solving a least-squares problem and using the quadratic formula. Here, Chua [4], starting with – and extending – ideas similar to ones above, devised and analyzed an algorithm for semidefinite programming (and, more generally, for symmetric cone programming) with complexity bounds matching the best presently known – $O(\sqrt{n} \log(1/\epsilon))$ iterations to reduce the duality gap by a factor ϵ when Λ_+ is the cone of $n \times n$ sdp matrices.

Although in the present work we do not analyze methods for efficiently computing $x_e^{(i)}$, we now present a few results relevant to algorithm design. These results also are important to the proof of the Main Theorem.

For the next two theorems, nothing is gained by distinguishing $p_e^{(i)}$ from any other hyperbolic polynomial, so we phrase the results simply in terms of a polynomial p hyperbolic in direction e , and its first derivative p'_e . (The two theorems, moreover, do not require Λ_+ to be a regular cone.)

Let

$$q_e := p/p'_e,$$

a rational function. The natural domain for q_e is Λ'_{++e} , because $p'_e(x) > 0$ for $x \in \Lambda'_{++e}$ and $p'_e(x) = 0$ for x in the boundary of Λ'_{++e} .

The following result is proven in §7.

Theorem 7. *The function $q_e : \Lambda'_{++e} \rightarrow \mathbb{R}$ is concave.*

Previously, q_e was known to be concave on the smaller cone Λ_{++} (see [1]). For us, the significance of the function being concave on the larger cone Λ'_{++} is that, as the following theorem illustrates, HP can be reformulated as a linearly-constrained convex optimization problem (no explicit conic constraint). For motivation, think of the situation where one has an approximation to an optimal solution for HP and the goal is to compute a better approximation.

Theorem 8. *The optimal solutions $x \in \Lambda'_{++e}$ for HP are the same as for the convex optimization problem*

$$\begin{aligned} \min_{x \in \Lambda'_{++e}} \quad & -\ln c^*(e - x) - q_e(x) \\ \text{s.t.} \quad & Ax = b. \end{aligned}$$

(A caution: The theorem asserts nothing about optimal solutions for HP that happen to lie in the intersection of the boundaries $\partial\Lambda_+$ and $\partial\Lambda'_{++e}$.) The theorem is proven in §7.

Since, by (2.3), $x_e^{(i)} \in \Lambda'_{++e}$ for $e \in \text{Swath}(i) \setminus \text{Core}(i)$, the following corollary is immediate, except for the assertion regarding the second differential, which is established in §7.

Corollary 9. *If $1 \leq i \leq n - 2$, $e \in \text{Swath}(i) \setminus \text{Core}(i)$ and*

$$f(x) := -\ln c^*(e - x) - \frac{p_e^{(i)}(x)}{p_e^{(i+1)}(x)},$$

then $x_e^{(i)}$ is the unique optimal solution for the convex optimization problem

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & Ax = b. \end{aligned}$$

Moreover, $D^2f(x_e^{(i)})[v, v] > 0$ if v is not a scalar multiple of $x_e^{(i)}$ (in particular, if $v \neq 0$ satisfies $Av = 0$).

A consequence of the assertion regarding the second differential is that Newton's method will converge quadratically to $x_e^{(i)}$ if initiated nearby. (Which is not to say that Newton's method is the algorithm of choice for this problem.)



Our final result relates the dynamics $\dot{e}(t) = x_{e(t)}^{(i)} - e(t)$ to the optimization problem dual to HP:

$$\left. \begin{array}{ll} \sup_{y^*, s^*} & y^* b \\ \text{s.t.} & y^* A + s^* = c^* \\ & s^* \in \Lambda_+^* \end{array} \right\} \text{HP}^*,$$

where Λ_+^* is the cone dual to Λ_+ .⁴ A pair (y^*, s^*) satisfying the constraints is said to be “strictly” feasible if $s^* \in \text{int}(\Lambda_+^*)$ (interior).

Letting val^* denote the optimal value of HP^* ($\text{val}^* = -\infty$ if HP^* is infeasible), we have, just as a matter of tracing definitions, the standard result known as “weak duality”⁵: $\text{val}^* \leq \text{val}$.

Optimizers expect that if a dynamics provides a natural path-following framework for solving a convex optimization problem, then not only do the dynamics generate paths leading to (primal) optimality, also the dynamics somehow generate paths leading to dual optimality.

For $e \in \text{Swath}(i) \setminus \text{Core}(i)$, define

$$s_e^{(i)} := \frac{c^*(e - x)}{p_e^{(i+1)}(x)} Dp_e^{(i)}(x) \quad \text{where } x = x_e^{(i)},$$

and where $Dp_e^{(i)}(x)$ is the differential of $p_e^{(i)}$ at x , i.e., the linear functional defined on vectors v by $Dp_e^{(i)}(x)[v] := \frac{d}{dt} p_e^{(i)}(x + tv)|_{t=0}$. The following result is proven in §9.

Main Theorem (Part II). *Assume $1 \leq i \leq n-2$ and let $t \mapsto e(t)$ ($0 \leq t < T$) be a maximal trajectory for the dynamics $\dot{e} = x_e^{(i)} - e$ starting at $e(0) \in \text{Swath}(i) \setminus \text{Core}(i)$. Then $y^* A + s_{e(t)}^{(i)} = c^*$ has a unique solution $y^* = y_{e(t)}^{(i)}$, and the pair $(y_{e(t)}^{(i)}, s_{e(t)}^{(i)})$ is strictly feasible for HP^* . Moreover,*

$$y_{e(t)}^{(i)} b = c^* x_{e(t)}^{(i)} \xrightarrow[t \rightarrow T]{} \text{val}$$

(in fact, increases to val strictly monotonically) and the path $t \mapsto (y_{e(t)}^{(i)}, s_{e(t)}^{(i)})$ is bounded.

Consequences are, of course, that $\text{val}^* = \text{val}$ (“strong duality”) and that the limit points of $t \mapsto (y_{e(t)}^{(i)}, s_{e(t)}^{(i)})$ form a nonempty set, each of whose elements is optimal for HP^* .

Thus, although the path $t \mapsto x_{e(t)}^{(i)}$ is infeasible for HP, and the feasible trajectory $t \mapsto e(t)$ can potentially not converge to optimality (it might instead converge to $\bar{e} \in \text{Core}(i)$), there is naturally generated a path $t \mapsto (y_{e(t)}^{(i)}, s_{e(t)}^{(i)})$ that both is feasible for HP^* and converges to optimality.

⁴The dual cone Λ_+^* consists of the linear functionals s^* satisfying $s^* x \geq 0$ for all $x \in \Lambda_+$.

⁵This is proven simply by observing that for feasible x and (y^*, s^*) ,

$$c^* x = (y^* A + s^*) x = y^* b + s^* x \geq y^* b,$$

the inequality due to $x \in \Lambda_+$ and $s^* \in \Lambda_+^*$.

Perhaps, then, the algorithmic framework we have posed as being for the primal optimization problem would be better posed as being for the dual, since by interchanging the primal and the dual, there would be a single path generated for the primal, and that path would be both feasible and converge to optimality.



Lastly, we mention that in defining the sequence of derivative polynomials $p_e^{(i)}$ ($i = 1, \dots, n-1$), we could have used various derivative directions e_1, \dots, e_{n-1} , choosing e_1 from the hyperbolicity cone for p and defining $p'_{e_1}(x) := Dp(x)[e_1]$, choosing e_2 from the (larger) hyperbolicity cone for p'_{e_1} and defining $p''_{e_1, e_2}(x) := Dp'_{e_1}(x)[e_2] = D^2p(x)[e_1, e_2]$, and so on. Several results in the following pages can be extended to this more general setting. However, computing multidirectional derivatives can be prohibitively expensive, even for the innocuous-appearing hyperbolic polynomials $p(x) = \prod_{i=1}^n a_i^T x$ naturally arising from polyhedral cones $\{x \in \mathbb{R}^n : a_i^T x \geq 0 \text{ for all } i = 1, \dots, n\}$; indeed, choosing e_1, \dots, e_n to be the standard basis, $D^n p(x)[e_1, \dots, e_n]$ is the permanent of the matrix whose i^{th} column is a_i .

By contrast, if the same direction e is used for all derivatives ($i = 1, \dots, n-1$), the resulting polynomials $p_e^{(i)}$ (resp., their gradients, their Hessians) can be efficiently evaluated at any point if the initial polynomial p (resp., its gradient, its Hessian) can be efficiently evaluated at any point. This is straightforwardly accomplished by interpolation, and can be sped up via the (inverse) Discrete Fourier Transform (see §9 of [13] for some discussion). As a primary motivation for the present paper is designing efficient algorithms, it thus is sensible to restrict consideration to the same direction e being used for all derivatives ($i = 1, \dots, n-1$).

3. Prelude to the Analysis

Now we turn to proving the results. The theorems are proven in the order in which they were stated with the exceptions of Part I of the Main Theorem and Theorem 6. The proof of the first is delayed because it depends on theorems that were stated later. The proof of Theorem 6 is delayed, until the end, due to the combination of the proof being long and the theorem being less important than others.

To ease burdens on the reader, each theorem is restated before its proof, and is renumbered to match the section in which it is proven (in part so it is clear there is no circularity among the proofs). Additionally, concepts and definitions are recalled as they first retake center stage, and a few supplementary results are presented. Thus, the reader is freed from having to refer to the preceding “overview of results.”

Several proofs rely fundamentally on results from [13], a paper on structural aspects of hyperbolicity cones (and hyperbolic programs), a paper for which a primary goal was to provide a ready reference of “lower-level” details so that subsequent papers (such as the present one) could avoid drawn-out proofs. The relevant results are presented in propositions or theorems at the beginning of sections where the results are first needed, with the exception of the following results.

One result from [13] is used time and time again – the characterization we recorded as (2.1), that is,

$$\Lambda_{+,e}^{(i)} = \{x : p_e^{(j)}(x) \geq 0 \text{ for all } j = i, \dots, n-1\}. \quad (3.1)$$

For ease of reference we record the following two characterizations that are similar to the one above:

$$\Lambda_{++,e}^{(i)} = \{x : p_e^{(j)}(x) > 0 \text{ for all } j = i, \dots, n-1\}, \quad (3.2)$$

$$\Lambda_{+,e}^{(i)} = \{x : p_e^{(i)}(x) > 0 \text{ and } p_e^{(j)}(x) \geq 0 \text{ for all } j = i+1, \dots, n-1\}. \quad (3.3)$$

As was noted earlier, (3.1) is immediate from Proposition 18 and Theorem 20 in [13]. The representations (3.2) and (3.3) are established in the two paragraphs following that theorem.

Readily proven from the above characterizations are that for $1 \leq i \leq n-1$,

$$x \in \Lambda_{+,e}^{(i)} \setminus \Lambda_{+,e}^{(i-1)} \Rightarrow p_e^{(i-1)}(x) < 0 \quad (3.4)$$

and

$$x \in \partial\Lambda_{+,e}^{(i)} \Rightarrow p_e^{(i-1)}(x) \leq 0. \quad (3.5)$$

On a different note, recall that the k^{th} differential $D^k f(x)$ of a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is defined on k -tuples of vectors by

$$D^k f(x)[v_1, \dots, v_k] := \frac{d}{dt_1} \cdots \frac{d}{dt_k} f(x + t_1 v_1 + \cdots + t_k v_k) \big|_{t_k=0} \cdots \big|_{t_1=0}.$$

If f is analytic – as are hyperbolic polynomials – the differential is well-defined for every k , and is symmetric (i.e., for all permutations π of $\{1, \dots, k\}$, the values $D^k f(x)[v_{\pi(1)}, \dots, v_{\pi(k)}]$ are identical). Moreover, the differentials are multilinear, that is, $v_i \mapsto D^k f(x)[v_1, \dots, v_k]$ is linear when $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k$ are fixed.

Fixing vectors w_1, \dots, w_j results in a symmetric, $(k-j)$ -multilinear form

$$[v_1, \dots, v_{k-j}] \mapsto D^k f(x)[w_1, \dots, w_j, v_1, \dots, v_{k-j}].$$

This form is denoted $D^k f(x)[w_1, \dots, w_j]$.

A fact used extensively, and which is easily proven (by, say, induction on j with base case $j = 0$), occurs when f is homogeneous of degree ℓ (i.e., $f(tx) = t^\ell f(x)$): For $0 \leq j \leq k \leq \ell$,

$$D^k f(x) \underbrace{[x, \dots, x]}_{j \text{ times}} = \frac{(\ell-k+j)!}{(\ell-k)!} D^{k-j} f(x).$$

In particular, for $0 \leq j \leq k \leq n-i$,

$$D^k p_e^{(i)}(x) \underbrace{[x, \dots, x]}_{j \text{ times}} = \frac{(n-i-k+j)!}{(n-i-k)!} D^{k-j} p_e^{(i)}(x). \quad (3.6)$$

A consequence of (3.6) used occasionally is that if $0 \leq k \leq n-i$, then

$$D^k p_e^{(i)}(e) = D^{k+i} p(e) \underbrace{[e, \dots, e]}_{i \text{ times}} = \frac{(n-k)!}{(n-i-k)!} D^k p(e) \quad (3.7)$$

– in particular, $D^k p_e^{(i)}(e)$ is a positive multiple of $D^k p(e)$.

Finally, we mention that the nesting of the derivative cones often is used implicitly. Two examples of assertions made without the nesting being mentioned: “If $x \in \text{Opt}_e^{(i)}$ and $x \in \text{Feas}$, then $x \in \text{Opt}$.” “If $x \in \Lambda_{+,e}^{(i)}$ then $p_e^{(i+1)}(x) > 0$.”

4. Proofs of Theorems 1 and 2, and the Nesting of Swaths

The following theorem records results from [13] that are used in this section.

Theorem 4.1. *For any polynomial hyperbolic in direction e , the following hold:*

- (A) For $i = 1, \dots, n-1$, $\Lambda_{+,e}^{(i-1)} \cap \partial\Lambda_{+,e}^{(i)} = \Lambda_+ \cap \partial\Lambda_{+,e}^{(i)}$.
- (B) For $i = 0, \dots, n-1$, the intersection $\Lambda_+ \cap \partial\Lambda_{+,e}^{(i)}$ is independent of $e \in \Lambda_{++}$.
- (C) For $i = 0, \dots, n-2$, the lineality space of $\Lambda_{+,e}^{(i)}$ consists precisely of the points in $\Lambda_{+,e}^{(i)} \cap \partial\Lambda_{+,e}^{(n-1)}$.
- (D) The cones $\Lambda_{+,e}^{(0)}, \Lambda_{+,e}^{(1)}, \dots, \Lambda_{+,e}^{(n-2)}$ have the same lineality space (thus, if one of the cones is regular, all are regular).

- (E) If Λ_+ is regular and $0 \leq i \leq n-2$, then any face of $\Lambda_{+,e}^{(i)}$ either is a face of Λ_+ or is an extreme direction of $\Lambda_{+,e}^{(i)}$.

Proof. Result (A) is Proposition 16 in [13]; (B) is established from Theorem 12 by induction (with base case $i = 0$) and by use of Proposition 22 (which shows that although the definition of the sets “ $\partial^m \Lambda_+$ ” appearing in Theorem 12 depends on a derivative direction e , the sets actually are independent of $e \in \Lambda_{++}$); (C) follows from Proposition 11; (D) is immediate from Proposition 13 by induction with base case $i = 1$ (alternatively, follows from (A) and (C)); (E) is Proposition 24. \square

Here again is Theorem 1, but renamed:

Corollary 4.2. Assume Λ_+ is regular and $0 \leq i \leq n-2$.

- (A) The intersection $\Lambda_+ \cap \partial \Lambda_{+,e}^{(i)}$ is independent of $e \in \Lambda_{++}$ (thus, a face of Λ_+ which is a boundary face of $\Lambda_{+,e}^{(i)}$ for some $e \in \Lambda_{++}$ is a boundary face for all $e \in \Lambda_{++}$).
- (B) If $e \in \Lambda_{++}$ then any boundary face of $\Lambda_{+,e}^{(i)}$ either is a face of Λ_+ or is a single ray contained in $\Lambda_{+,e}^{(i+1)} \setminus \Lambda_{+,e}^{(i-1)}$.

Proof. Parts (B) and (E) of Theorem 4.1 mostly prove the corollary. It remains only to show

$$x \in (\partial \Lambda_{+,e}^{(i)}) \setminus \Lambda_+ \Rightarrow x \in \Lambda_{+,e}^{(i+1)} \setminus \Lambda_{+,e}^{(i-1)}. \quad (4.1)$$

However, Theorem 4.1(A) immediately gives

$$x \in (\partial \Lambda_{+,e}^{(i)}) \setminus \Lambda_+ \Rightarrow x \notin \Lambda_{+,e}^{(i-1)},$$

and upon substituting $i+1$ for i , immediately gives

$$x \in \Lambda_{+,e}^{(i)} \setminus \Lambda_+ \Rightarrow x \notin \partial \Lambda_{+,e}^{(i+1)},$$

that is, gives

$$x \in \Lambda_{+,e}^{(i)} \setminus \Lambda_+ \Rightarrow x \in \Lambda_{+,e}^{(i+1)}.$$

The implication (4.1) is thus established, concluding the proof. \square

Throughout the remainder of the paper, our standard assumptions apply without being made explicit in the statements of theorems, propositions, etc. Recall these assumptions regard the objective function c^*x , the equations $Ax = b$, and the cone Λ_+ :

- $b \neq 0$ (in particular, the origin is infeasible)
- A is surjective (i.e., onto)
- c^* is not in the image of A^* (otherwise all feasible points would be optimal)
- Λ_+ is a regular cone, that is, contains no subspaces other than $\{0\}$

Thus, by Theorem 4.1(D), $\Lambda_{+,e}^{(i)}$ is regular for all $0 \leq i \leq n-2$.

Recall the definitions of swaths and cores:

$$\text{Swath}(i) := \{e \in \Lambda_{++} : Ae = b \text{ and } \text{Opt}_e^{(i)} \neq \emptyset\},$$

$$\text{Core}(i) := \{e \in \text{Swath}(i) : \text{Opt}_e^{(i)} = \text{Opt}\}.$$

Recall, too, that the above corollary (aka Theorem 1) and our standard assumptions easily imply that if $1 \leq i \leq n-2$ and $e \in \text{Swath}(i) \setminus \text{Core}(i)$, then $\text{Opt}_e^{(i)}$ consists of a single point, denoted $x_e^{(i)}$. The corollary implies, moreover, that

$$x_e^{(i)} \in \text{relint}(\text{Feas}_e^{(i+1)}) \setminus \text{Feas}_e^{(i-1)}. \quad (4.2)$$

In §2 we noted that from the nesting

$$\text{Feas}_e^{(0)} \subseteq \text{Feas}_e^{(1)} \subseteq \dots \subseteq \text{Feas}_e^{(n-1)}$$

easily follows the nesting of cores,

$$\text{Core}(0) \supseteq \text{Core}(1) \supseteq \cdots \supseteq \text{Core}(n-1).$$

We now establish (a bit more than) the nesting of swaths.

Proposition 4.3. *The swaths are nested,*

$$\text{Swath}(0) \supseteq \text{Swath}(1) \supseteq \cdots \supseteq \text{Swath}(n-1)$$

(thus, if any swath is nonempty, so is $\text{Swath}(0)$ – equivalently, so is Opt).

Moreover, if $\text{Swath}(n-1) \neq \emptyset$, or if $\text{Swath}(i) \setminus \text{Core}(i) \neq \emptyset$ for some $1 \leq i \leq n-2$, then Opt is a bounded set.

Proof. We know $\text{Core}(i) \subseteq \text{Core}(i-1)$, so to prove the nesting of swaths, we assume $e \in \text{Swath}(i) \setminus \text{Core}(i)$ and show $e \in \text{Swath}(i-1)$.

First consider the case $1 \leq i \leq n-2$, where $\text{Opt}_e^{(i)}$ consists of the single point $x_e^{(i)}$ – in particular, $\text{Opt}_e^{(i)}$ is nonempty and bounded, from which follows by standard convexity arguments that the level sets $\{x \in \text{Feas}_e^{(i)} : c^*x \leq \alpha\}$ ($\alpha \in \mathbb{R}$) are bounded. Since $\text{Feas}_e^{(i-1)} \subseteq \text{Feas}_e^{(i)}$, the level sets $\{x \in \text{Feas}_e^{(i-1)} : c^*x \leq \alpha\}$ also are bounded. From this and the nonemptiness of $\text{Feas}_e^{(i-1)}$ (indeed, $e \in \text{Feas}_e^{(i-1)}$) easily follows $\text{Opt}_e^{(i-1)} \neq \emptyset$, that is, $e \in \text{Swath}(i-1)$, as desired.

Observe, too, that the boundedness of the level sets $\{x \in \text{Feas}_e^{(i)} : c^*x \leq \alpha\}$, and the relations $\text{Feas}_e^{(i)} \supseteq \text{Feas}_e^{(j)}$ for $j = 0, \dots, i$, imply $\text{Opt}_e^{(j)}$ is bounded for all $j = 0, \dots, i$ – in particular, $\text{Opt} = \text{Opt}_e^{(0)}$ is bounded, thereby establishing the final statement of the proposition for the case that $\text{Swath}(i) \setminus \text{Core}(i) \neq \emptyset$ for some $1 \leq i \leq n-2$. (We note as an aside that always, $\text{Core}(0) = \text{Swath}(0)$, so never is $\text{Core}(0)$ a proper subset of $\text{Swath}(0)$.)

We have left to prove that $\text{Swath}(n-1) \subseteq \text{Swath}(n-2)$, and that if $\text{Swath}(n-1) \neq \emptyset$, then Opt is bounded. Here we rely on the following result, which is immediate from (C) and (D) of Theorem 4.1 :

(F) For $i = 0, \dots, n-2$, the lineality space of Λ_+ consists precisely of the points in $(\partial\Lambda_{+,e}^{(n-1)}) \cap \Lambda_{+,e}^{(i)}$.

Now, for every $e \in \Lambda_{++}$, the cone $\Lambda_{+,e}^{(n-1)}$ is a halfspace, i.e., $\Lambda_{+,e}^{(n-1)} = \{x : d^*x \geq 0\}$ for some linear functional d^* . Moreover, since Λ_+ is a regular cone, the hyperplane $\{x : d^*x = 0\}$ intersects the cone $\Lambda_{+,e}^{(n-2)}$ only at the origin (by (F) with $i = n-2$). Consequently, the level sets $\{x \in \Lambda_{+,e}^{(n-2)} : d^*x \leq \alpha\}$ ($\alpha \in \mathbb{R}$) are bounded.

Assume $e \in \text{Swath}(n-1)$. Then, clearly,

$$\text{Opt}_e^{(n-1)} = \{x : Ax = b \text{ and } d^*x = 0\},$$

and the first-order conditions are satisfied:

$$c^* = y^*A + \lambda d^* \quad \text{for some } \lambda \geq 0 \text{ and } y^*.$$

Since c^* is not in the image of A^* (by assumption), it must be that $\lambda > 0$, from which follows that each level set $\{x \in \Lambda_{+,e}^{(n-2)} : Ax = b \text{ and } c^*x \leq \beta\}$ is a level set $\{x : Ax = b \text{ and } d^*x \leq \alpha\}$ for some α (depending on β), and hence is bounded (according to the conclusion of the preceding paragraph). From this easily follows $\text{Opt}_e^{(n-2)} \neq \emptyset$, that is, $e \in \text{Swath}(n-2)$. It also follows, of course, that for all $0 \leq i \leq n-2$, the level sets $\{x \in \Lambda_{+,e}^{(i)} : Ax = b \text{ and } c^*x \leq \beta\}$ are bounded, and hence that $\text{Opt}_e^{(i)}$ is bounded – in particular, $\text{Opt} = \text{Opt}_e^{(0)}$ is bounded. \square

We close this section by restating and proving Theorem 2.

Theorem 4.4. $\text{Swath}(n-1) = \text{Central Path}$

Proof. The central path consists precisely of the points in $\{x \in \Lambda_{++} : Ax = b\}$ which minimize $f_\eta(x) := \eta c^*x - \ln p(x)$ for some $\eta > 0$. As the functions η are convex, the first-order optimality conditions are sufficient as well as necessary; thus, the central path consists precisely of the points e satisfying

$$\left. \begin{aligned} \eta c^* - \frac{1}{p(e)} Dp(e) &= y^* A \quad \text{for some } \eta > 0 \text{ and } y^* \\ Ae &= b \\ e &\in \Lambda_{++} . \end{aligned} \right\} \quad (4.3)$$

For every $e \in \Lambda_{++}$, on the other hand, from the linearity of $x \mapsto p_e^{(n-1)}(x)$, we have $p_e^{(n-1)}(x) = Dp_e^{(n-1)}(z)[x]$ for every point z – in particular for $z = e$ – and thus

$$\Lambda_{+,e}^{(n-1)} = \{x : Dp_e^{(n-1)}(e)[x] \geq 0\} .$$

Clearly, then, $\text{Opt}_e^{(n-1)} \neq \emptyset$ if and only if

$$c^* = \lambda Dp_e^{(n-1)}(e) + w^* A \quad \text{for some } \lambda \geq 0 \text{ and } w^* .$$

Since c^* is not in the range of A^* (by assumption), and since $Dp_e^{(n-1)}(e) = (n-1)! Dp(e)$ (by (3.7)), we thus have that $e \in \text{Swath}(n-1)$ if and only if

$$\left. \begin{aligned} c^* &= \lambda (n-1)! Dp(e) + w^* A \quad \text{for some } \lambda > 0 \text{ and } w^* \\ Ae &= b \\ e &\in \Lambda_{++} . \end{aligned} \right\} \quad (4.4)$$

Obviously, the conditions (4.4) and (4.3) are equivalent. \square

5. Proof of Theorem 3

Having finished proving that for each $e \in \text{Swath}(i) \setminus \text{Core}(i)$, the set $\text{Opt}_e^{(i)}$ consists of a unique point $x_e^{(i)}$, it is time to show that the differential equation

$$\frac{d}{dt} e(t) = x_{e(t)}^{(i)} - e(t), \quad e(0) \in \text{Swath}(i) \setminus \text{Core}(i)$$

results in well-defined trajectories. This is immediate from Theorem 3, which we now restate and prove.

Theorem 5.1. *Assume $1 \leq i \leq n-2$. The set $\text{Swath}(i) \setminus \text{Core}(i)$ is open in the relative topology of $\text{relint}(\text{Feas})$. Moreover, the map $e \mapsto x_e^{(i)}$ is analytic on $\text{Swath}(i) \setminus \text{Core}(i)$.*

To establish the theorem, we introduce another theorem and a proposition. The proof of the proposition is left to the reader, as it follows entirely standard lines, and is primarily an application of the Implicit Function Theorem to optimality conditions (see, for example, §2.4 of [5] for similar results).

Proposition 5.2. *Assume $f : \mathcal{E}_1 \times \mathcal{E}_2 \rightarrow \mathbb{R}$ is analytic, where \mathcal{E}_1 and \mathcal{E}_2 are Euclidean spaces. For $z \in \mathcal{E}_2$, define $f_z : \mathcal{E}_1 \rightarrow \mathbb{R}$ by $f_z(y) := f(y, z)$, and consider for some linear functional $y \mapsto d^*y$ the following family of optimization problems parameterized by z :*

$$\begin{aligned} \min_y \quad & d^*y \\ \text{s.t.} \quad & f_z(y) \geq 0 . \end{aligned} \quad (5.1)$$

For some \bar{z} , assume \bar{y} is a local optimum with the properties that $Df_{\bar{z}}(\bar{y}) \neq 0$, and $D^2 f_{\bar{z}}(\bar{y})[v, v] < 0$ for all $v \neq 0$ satisfying $Df_{\bar{z}}(\bar{y})[v] = 0$. Then there exists an analytic function $z \mapsto y_z$ defined on an open neighborhood of \bar{z} , and possessing the properties $y_{\bar{z}} = \bar{y}$ and for each z in the neighborhood, the point y_z is locally optimal for (5.1).

The following theorem collects results from [13]. (Keep in mind that, as was emphasized in §4, our standard assumptions are always assumed to be in effect. The relevant assumption below is regularity of Λ_+ , but this assumption is immaterial for parts (A) and (B).)

Theorem 5.3. *Assume $1 \leq i \leq n-1$, $\bar{e} \in \Lambda_{++}$ and $\bar{x} \in (\partial\Lambda_{+, \bar{e}}^{(i)}) \setminus \Lambda_{+, \bar{e}}^{(i-1)}$ (equivalently, by Theorem 4.1(A), $\bar{x} \in (\partial\Lambda_{+, \bar{e}}^{(i)}) \setminus \Lambda_+$).*

(A) *There exist open neighborhoods U of \bar{e} , and V of \bar{x} , with the property that*

$$e \in U \quad \Rightarrow \quad \Lambda_{+, e}^{(i)} \cap V = \{x \in V : p_e^{(i)}(x) \geq 0\}.$$

(B) *In a neighborhood of \bar{x} , $\partial\Lambda_{+, \bar{e}}^{(i)}$ is a manifold, whose tangent space at \bar{x} is $\{v : Dp_{\bar{e}}^{(i)}(\bar{x})[v] = 0\}$.*

(C) *Further assume $i \leq n-2$. Then $D^2p_{\bar{e}}^{(i)}(\bar{x})[v, v] < 0$ for all vectors v that both satisfy $Dp_{\bar{e}}^{(i)}(\bar{x})[v] = 0$ and are not scalar multiples of \bar{x} .*

Proof. Recall the characterization (3.1), that is,

$$\Lambda_{+, e}^{(i)} = \{x : p_e^{(j)}(x) \geq 0 \text{ for all } j = i, \dots, n-1\}. \quad (5.2)$$

For $i = n-1$, result (B) is trivial because $x \mapsto p_e^{(n-1)}(x)$ is a linear functional, and (A) is immediate with the additional observation that the linear functional varies continuously in e . Henceforth we assume $1 \leq i \leq n-2$.

Theorem 4.1(A) gives

$$\bar{x} \in (\partial\Lambda_{+, \bar{e}}^{(i)}) \setminus \Lambda_+ \quad \Rightarrow \quad \bar{x} \in \Lambda_{+, \bar{e}}^{(i+1)},$$

and hence,

$$p_{\bar{e}}^{(j)}(\bar{x}) > 0 \quad \text{for } j = i+1, \dots, n-1. \quad (5.3)$$

Thus, since the polynomials vary continuously in e as well as in x , there exist open neighborhoods U of \bar{e} , and V of \bar{x} , for which

$$(e, x) \in U \times V \quad \Rightarrow \quad p_e^{(j)}(x) > 0 \text{ for all } j = i+1, \dots, n-1.$$

This and (5.2) establish statement (A) of the present theorem.

In light of (A), to establish (B) we need only show $Dp_{\bar{e}}^{(i)}(\bar{x}) \neq 0$. However, $Dp_{\bar{e}}^{(i)}(\bar{x})[\bar{e}] = p_{\bar{e}}^{(i+1)}(\bar{x}) > 0$ (by (5.3); hence, $Dp_{\bar{e}}^{(i)}(\bar{x}) \neq 0$.

Finally, (C) is a restatement of Theorem 14 in [13], with $\Lambda_{+, e}^{(i-1)}$ substituted for Λ_+ (the hypothesis of regularity is satisfied due to Theorem 4.1(D)). \square

Proof of Theorem 5.1. Assume $1 \leq i \leq n-2$ and $e \in \text{Swath}(i) \setminus \text{Core}(i)$.

We claim

$$Dp_e^{(i)}(x_e^{(i)})[e - x_e^{(i)}] \neq 0. \quad (5.4)$$

Indeed, as $x_e^{(i)} \in (\partial\Lambda_{+, e}^{(i)}) \setminus \Lambda_{+, e}^{(i-1)}$ (using (4.2), the tangent space of $\Lambda_{+, e}^{(i)}$ at $x_e^{(i)}$ is precisely $\{v : Dp_e^{(i)}(x_e^{(i)})[v] = 0\}$, by Theorem 5.3(B). However, e is in the interior of the convex cone $\Lambda_{+, e}^{(i)}$, and hence the difference $e - x_e^{(i)}$ cannot be in the tangent space. The claim is thus established.

Let $\bar{\mathcal{E}} := \{x : Ax = 0\}$, a Euclidean (sub)space. Define $f : \bar{\mathcal{E}} \times \bar{\mathcal{E}} \rightarrow \mathbb{R}$ by $f(y, z) = p_{z+e}^{(i)}(y + e)$, and let $f_z : \bar{\mathcal{E}} \rightarrow \mathbb{R}$ be the function $f_z(y) := f(y, z)$.

Let d^* be the projection of c^* onto $\bar{\mathcal{E}}$ (i.e., the linear functional on $\bar{\mathcal{E}}$ satisfying $d^*y = c^*y$ for all $y \in \bar{\mathcal{E}}$). For any x satisfying $Ax = b$, the projection of $Dp_e^{(i)}(x)$ onto $\bar{\mathcal{E}}$ is $Df_0(y)$ where $y = x - e$. In particular, for all $v \in \bar{\mathcal{E}}$,

$$Dp_e^{(i)}(x_e^{(i)})[v] = Df_0(\bar{y})[v] \quad \text{where } \bar{y} := x_e^{(i)} - e.$$

Since $e - x_e^{(i)} \in \bar{\mathcal{E}}$, (5.4) thus implies $Df_0(\bar{y}) \neq 0$.

Similarly, for all $v \in \bar{\mathcal{E}}$,

$$D^2 f_0(\bar{y})[v, v] = D^2 p_e^{(i)}(x_e^{(i)})[v, v] .$$

Thus, since $x_e^{(i)} \notin \bar{\mathcal{E}}$ (because $b \neq 0$, by assumption), Theorem 5.3(C) implies $D^2 f_0(\bar{y})[v, v] < 0$ for all $0 \neq v \in \bar{\mathcal{E}}$ satisfying $D f_0(\bar{y})[v] = 0$.

In all, the hypotheses of Proposition 5.2 are satisfied for $\bar{z} = 0$ and \bar{y} . Letting $z \mapsto y_z$ be the analytic function whose existence is ensured by the proposition, it is readily apparent that $y_z + e$ is locally optimal for

$$\begin{aligned} \min_x \quad & c^* x \\ \text{s.t.} \quad & Ax = b \\ & p_{z+e}^{(i)}(x) \geq 0 . \end{aligned}$$

Since $z \mapsto y_z$ is continuous (it's even analytic), it thus follows from Theorem 5.3(A) that for z in an open neighborhood of the origin in $\bar{\mathcal{E}}$, the point $y_z + e$ is locally optimal – and hence globally optimal – for the convex optimization problem $\text{HP}_{z+e}^{(i)}$ – that is, $y_z + e \in \text{Opt}_{z+e}^{(i)}$, and so $z + e \in \text{Swath}(i)$.

However, for z in a possibly smaller open neighborhood of the origin, $y_z + e \notin \Lambda_+$, because $y_0 + e = x_e^{(i)} \notin \Lambda_+$ and $z \mapsto y_z$ is continuous. Consequently, for z in this open neighborhood of the origin, $\text{Opt}_{z+e}^{(i)} \setminus \Lambda_+ \neq \emptyset$, and hence, $e \notin \text{Core}(i)$; thus, $z \in \text{Swath}(i) \setminus \text{Core}(i)$, and clearly, $x_{z+e}^{(i)} = y_z + e$.

As e was an arbitrary point in $\text{Swath}(i) \setminus \text{Core}(i)$, the proof is complete. \square

In closing this section, we recall that due simply to the strict curvature of (regular) second order cones, for every $e \in \text{Swath}(n-2)$ there is a unique optimal solution of $\text{HP}_e^{(n-2)}$, thus naturally extending the map $e \mapsto x_e^{(n-2)}$ to all of $\text{Swath}(n-2)$. Additionally, every $\bar{e} \in \Lambda_{++}$ and every $0 \neq x \in \partial \Lambda_{+, \bar{e}}^{(n-2)}$ satisfies properties (A), (B) and (C) of Theorem 5.3 when $i = n-2$ (regardless of whether \bar{x} satisfies the theorem's hypothesis $\bar{x} \notin \Lambda_+$). Consequently, the proof of Theorem 5.1 easily is made to be a proof showing that the extended map $e \mapsto x_e^{(n-2)}$ is analytic on $\text{Swath}(n-2)$ (and $\text{Swath}(n-2)$ is open in the relative topology of $\text{relint}(\text{Feas})$).

6. Proof of Theorem 5

We have been proving the theorems of §2 in the order they appeared, and have just finished establishing Theorem 3, that is, have just finished establishing the well-definedness of the trajectories arising from the differential equation

$$\frac{d}{dt} e(t) = x_{e(t)}^{(i)} - e(t), \quad e(0) \in \text{Swath}(i) \setminus \text{Core}(i) .$$

To continue proving theorems in the order they appeared, we would next prove Part I of the Main Theorem, which states that either the trajectory $t \mapsto e(t)$ or the path $t \mapsto x_{e(t)}^{(i)}$ converges to optimality (perhaps both). However, there is groundwork to be laid before proving that result, so we jump to the one stated after it, a motivational theorem pertaining to the trajectories in the special case $i = n-2$.

For any e , the polynomial $x \mapsto p_e^{(n-2)}(x)$ is quadratic, and hence for $e \in \Lambda_{++}$, $\Lambda_{+, e}^{(n-2)}$ is a second-order cone, regularity due to Theorem 4.1(D).

Recall we say that $t \mapsto e(t)$ ($0 \leq t < T$) is a “maximal trajectory” for the dynamics $\dot{e}(t) = x_{e(t)}^{(i)} - e(t)$ if $e(0) \in \text{Swath}(i) \setminus \text{Core}(i)$ and T ($= T(e(0))$) is the time when the trajectory either reaches the boundary of $\text{Swath}(i) \setminus \text{Core}(i)$ or escapes to infinity.

Following is a restatement and proof of Theorem 5, whose only purposes in §2 were to motivate the choice of terminology “central swaths” and to give insight into

the origin of the idea that in general the trajectories $t \mapsto e(t)$ (or paths $t \mapsto x_{e(t)}^{(i)}$) converge to optimality. After completing the paper, the reader likely will not have much difficulty in making the statement of the theorem more complete.⁶

Theorem 6.1. *Assume $t \mapsto e(t)$ ($0 \leq t < T$) is a maximal trajectory arising from the dynamics $\dot{e}(t) = x_{e(t)}^{(n-2)} - e(t)$, starting at $e(0) \in \text{Swath}(n-2) \setminus \text{Core}(n-2)$.*

If $e(0) \in \text{Central Path}$ then $\{e(t) : 0 \leq t < T\} \subseteq \text{Central Path}$.

Proof. The proof makes use of results known to anyone familiar with the interior-point method literature.

The central path is the set $\{z(\eta) : \eta > 0\}$, where $z(\eta)$ solves

$$\begin{aligned} \min_x \quad & \eta c^* x - \ln p(x) \\ \text{s.t.} \quad & Ax = b \\ & x \in \Lambda_{++}. \end{aligned}$$

In the terminology of Nesterov and Nemirovski [12], the barrier function $x \mapsto -\ln p(x)$ is self-concordant, a fact whose implications were first developed by Güler [7].

The barrier function is strictly convex on its domain, Λ_{++} (“strictly” because Λ_+ is a regular cone). Hence, for each $\eta > 0$, there is at most one optimizer, $z(\eta)$. It is well-known from interior-point method theory that if the optimizer $z(\eta)$ exists for one positive value of η , then an optimizer $z(\eta)$ exists for all $\eta > 0$, and the path $\eta \mapsto z(\eta)$ is analytic (using that $x \mapsto -\ln p(x)$ is analytic).

To prove the theorem, it suffices to show that if $z(\eta)$ lies in $\text{Swath}(i) \setminus \text{Core}(i)$ (which we know by Theorem 5.1 to be open in the relative topology of $\text{relint}(\text{Feas})$), then $\dot{z}(\eta)$ – the tangent vector to the central path at $z(\eta)$ – is a positive multiple of the difference $x_{z(\eta)}^{(n-2)} - z(\eta)$. Indeed, this is sufficient to ensure that if $z(\eta_0) \in \text{Swath}(i) \setminus \text{Core}(i)$, then as η increases from η_0 , the central path will remain in $\{e(t) : 0 \leq t < T(z(\eta_0))\}$ – where $e(0) = z(\eta_0)$ – until reaching the boundary of $\text{Swath}(i) \setminus \text{Core}(i)$, that is, will remain in the path of the trajectory until the trajectory ends, implying $\{e(t) : 0 \leq t < T(z(\eta_0))\} \subseteq \text{Central Path}$.

Now we begin proving that $\dot{z}(\eta)$ is indeed a positive multiple of $x_{z(\eta)}^{(n-2)} - z(\eta)$.

Assuming the central path exists, for each $\eta > 0$ there exists $y(\eta)^*$ which together with $z(\eta)$ satisfies the first-order optimality condition

$$\eta c^* = \frac{1}{p(z(\eta))} Dp(z(\eta)) + y(\eta)^* A. \quad (6.1)$$

It is well known that if A is surjective (one of our standard assumptions), then $y(\eta)^*$ is unique, and $\eta \mapsto y(\eta)^*$ is analytic (using that $x \mapsto -\ln p(x)$ is analytic).

⁶More specifically, after understanding the paper the reader likely will not have much difficulty in showing the following, in which

$$\text{Central Path} = \{z(\eta) : \eta > 0\},$$

$$\text{ECentral Path} = \{z(\eta) : \eta \in \mathbb{R}\}$$

(“extended central path”), and $z(\eta)$ solves $\min_x \eta c^* x - \ln p(x)$, s.t., $Ax = b$, $x \in \Lambda_{++}$:

- $\text{Central Path} \subseteq \text{Swath}(n-2)$
- If $e(0) \in \text{ECentral Path} \cap (\text{Swath}(n-2) \setminus \text{Core}(n-2))$, then $T(e(0)) = \infty$ and $\{e(t) : 0 \leq t < \infty\} = \{e \in \text{ECentral Path} : c^* e \leq c^* e(0)\}$.
- If $e \in \text{ECentral Path} \cap \text{Core}(n-2)$ then

$$\text{Core}(n-2) = \ell \cap \Lambda_{++} = \text{ECentral Path},$$

where ℓ is the line containing e and the (unique) point in $\text{Opt}_e^{(n-2)} = \text{Opt}$.

Differentiating in η provides equations satisfied by $\dot{z}(\eta)$, the tangent vector to the central path:

$$c^* = \frac{1}{p(z(\eta))} D^2 p(z(\eta))[\dot{z}(\eta)] - \frac{Dp(z(\eta))[\dot{z}(\eta)]}{p(z(\eta))^2} Dp(z(\eta)) + \dot{y}(\eta)^* A. \quad (6.2)$$

It is well known, moreover, that $c^* \dot{z}(\eta) < 0$.

Using (6.1) to substitute for $Dp(z(\eta))$ in (6.2), and relying on $A\dot{z}(\eta) = 0$, yields

$$\gamma(\eta)c^* = D^2 p(z(\eta))[\dot{z}(\eta)] + w(\eta)^* A \quad (6.3)$$

for some constant $\gamma(\eta)$ and vector $w(\eta)^*$. We aim to use this condition to show that if $z(\eta) \in \text{Swath}(n-2) \setminus \text{Core}(n-2)$, then $\dot{z}(\eta)$ is a positive multiple of $x_{z(\eta)}^{(n-2)} - z(\eta)$. To accomplish this aim, we need an appropriate characterization of $x_{z(\eta)}^{(n-2)}$, the optimal solution to $\text{HP}_{z(\eta)}^{(n-2)}$.

Now, for every $e \in \Lambda_{++}$, it holds that $\Lambda_{+,e}^{(n-2)} \cap \partial\Lambda_{+,e}^{(n-1)} = \{0\}$ (by regularity, and parts C and D of Theorem 4.1). Thus, $\text{Feas}_e^{(n-2)} \cap \partial\Lambda_{+,e}^{(n-1)} = \emptyset$ (because $b \neq 0$) and, as a consequence,

$$x \in \text{Feas}_e^{(n-2)} \Rightarrow p_e^{(n-1)}(x) > 0 \quad (6.4)$$

$$\begin{aligned} &\Leftrightarrow Dp_e^{(n-2)}(x)[e] > 0 \\ &\Rightarrow Dp_e^{(n-2)}(x) \neq 0. \end{aligned} \quad (6.5)$$

On the other hand, by the characterization (3.1) applied to $\Lambda_{+,e}^{(n-2)}$,

$$\text{Feas}_e^{(n-2)} = \{x : Ax = b, p_e^{(n-2)}(x) \geq 0 \text{ and } p_e^{(n-1)}(x) \geq 0\}. \quad (6.6)$$

It follows from (6.4), (6.5) and (6.6) that necessary and sufficient conditions for a point x to be optimal for the convex optimization problem $\text{HP}_e^{(n-2)}$ are

$$\left. \begin{aligned} \gamma c^* &= Dp_e^{(n-2)}(x) + y^* A \quad \text{for some } \gamma > 0 \text{ and } y^* \\ Ax &= b \\ x &\in \partial\Lambda_{+,e}^{(n-2)} \end{aligned} \right\} \quad (6.7)$$

Assume $z(\eta) \in \text{Swath}(i) \setminus \text{Core}(i)$. For brevity, write z for $z(\eta)$ and \dot{z} for $\dot{z}(\eta)$. Our goal is to show \dot{z} is a positive multiple of $x_z^{(n-2)} - z$ – equivalently, to show $x_z^{(n-2)} = z + t\dot{z}$ for some $t > 0$ – equivalently, to show for some $t > 0$ that $x = z + t\dot{z}$ satisfies the conditions (6.7) when $e = z$.

To fix the value of t , consider the condition $x \in \partial\Lambda_{+,z}^{(n-2)}$. To see that there exists $t > 0$ for which $x = z + t\dot{z}$ is in the boundary, begin by recalling $c^* \dot{z} < 0$, and hence by (6.1), $Dp(z)[\dot{z}] < 0$. However, by (3.7), $Dp(z)$ is a positive multiple of $Dp_z^{(n-1)}(z)$ – thus, $Dp_z^{(n-1)}(z)[\dot{z}] < 0$. Since $x \mapsto p_z^{(n-1)}(x)$ is linear and $p_z^{(n-1)}(z) > 0$, it follows that for some $t > 0$, the point $z + t\dot{z}$ lies in $\partial\Lambda_{+,z}^{(n-1)}$. By the nesting of cones, for a (smaller) value $t > 0$, the point $z + t\dot{z}$ lies in $\partial\Lambda_{+,z}^{(n-2)}$. Consider the value of t to be fixed thusly.

Of course $x = z + t\dot{z}$ satisfies $Ax = b$. To complete the proof, it remains to show there exist $\gamma > 0$ and y^* satisfying

$$\gamma c^* = Dp_z^{(n-2)}(z + t\dot{z}) + y^* A. \quad (6.8)$$

We claim, however, there is no need to be concerned with the sign of γ , because so long as (6.8) is satisfied, γ is forced to be positive. Indeed, if (6.8) held with $\gamma < 0$, then $z + t\dot{z}$ would satisfy the (sufficient as well as necessary) optimality conditions for the problem obtained from $\text{HP}_z^{(n-2)}$ by replacing “ $\min c^* x$ ” with “ $\max c^* x$ ”, contradicting that $c^* z > c^*(z + t\dot{z})$. And if (6.8) held with $\gamma = 0$, then by Theorem 5.3(B), $\{x : Ax = 0\}$ would be a subspace of the tangent space of $\partial\Lambda_{+,z}^{(n-2)}$ at $z + t\dot{z}$ – but this would lead to a contradiction, as z is in the interior of

$\Lambda_{+,z}^{(n-2)}$ and $A\dot{z} = 0$. Thus, we need only be concerned with showing there exist γ and y^* satisfying (6.8), and not be concerned with the sign of γ .

By linearity of $x \mapsto Dp_z^{(n-2)}(x)$ (as $p_z^{(n-2)}$ is a quadratic polynomial),

$$Dp_z^{(n-2)}(z + t\dot{z}) = Dp_z^{(n-2)}(z) + tDp_z^{(n-2)}(\dot{z}) .$$

Since $Dp_z^{(n-2)}(z)$ is a positive multiple of $Dp(z)$ (by (3.7)), and since $Dp(z)$ satisfies (6.1), to show that $x = e + t\dot{z}$ satisfies (6.8) for some γ and y^* , it thus suffices to show

$$Dp_e^{(n-2)}(\dot{z}) = \bar{\gamma}c^* + \bar{y}^*A \quad \text{for some } \bar{\gamma} \text{ and } \bar{y}^* . \quad (6.9)$$

Now, for any vectors u and v ,

$$\begin{aligned} D^2p_z^{(n-2)}(u)[v] &= \frac{d}{dt} \left(Dp_z^{(n-2)}(u + tv) \right) \Big|_{t=0} \\ &= Dp_z^{(n-2)}(v) \quad (\text{by linearity of } x \mapsto Dp_z^{(n-2)}(x)) . \end{aligned}$$

In particular,

$$Dp_z^{(n-2)}(\dot{z}) = D^2p_z^{(n-2)}(z)[\dot{z}] ,$$

and hence, by (3.7),

$$Dp_z^{(n-2)}(\dot{z}) \text{ is a positive multiple of } D^2p(z)[\dot{z}] .$$

Consequently, that $\bar{\gamma}$ and \bar{y}^* can be chosen to satisfy (6.8) follows from (6.3), thus completing the proof. \square

7. Proofs of Theorems 7 and 8

We now consider the optimization problems $\text{HP}_e^{(i)}$ for all $0 \leq i \leq n-1$. The purpose of the present section is to establish a useful characterization of those optimal solutions of $\text{HP}_e^{(i)}$ which do not lie in $\partial\Lambda_{+,e}^{(i+1)}$. For $e \in \text{Swath}(i) \setminus \text{Core}(i)$, the characterization precisely identifies the unique optimal solution $x_e^{(i)}$ (because $x_e^{(i)} \notin \partial\Lambda_{+,e}^{(i+1)}$ by (4.2)). Although only the case $e \in \text{Swath}(i) \setminus \text{Core}(i)$ is relevant to the Main Theorem, we record the characterization generally, because it has the potential for computational relevance also when $e \in \text{Core}(i)$.

In this section, the only properties used of the polynomials $p_e^{(i)}$ is that they are hyperbolic and nonconstant. **Thus, to ease notation, for this section we let p be any nonconstant polynomial which is hyperbolic in direction e .** We let p'_e denote the derivative polynomial.

If p is linear, then $p'_e \not\equiv 0$ is a constant polynomial and its hyperbolicity cone $\Lambda'_{+,e}$ is the entire Euclidean space \mathcal{E} .

To aid the reader's intuition, the following proposition (not used in the sequel) explains which points in Opt can possibly be not in $\partial\Lambda'_{+,e}$.

Proposition 7.1. *If $e \in \Lambda_{++}$ then*

$$\text{either } \text{Opt} \subset \partial\Lambda'_{+,e} \quad \text{or} \quad \text{Opt} \setminus \partial\Lambda'_{+,e} = \text{relint}(\text{Opt}) .$$

Proof. The proposition is trivially true when p is linear, because then, $\partial\Lambda'_{+,e} = \emptyset$ and $\text{relint}(\text{Opt}) = \text{Opt}$ (due to Opt being an affine space). The proposition also is trivially true if Opt is the empty set or consists of a single point. Thus, assume $\deg(p) \geq 2$, assume Opt has more than one point, and assume Opt is not contained in $\partial\Lambda'_{+,e}$.

Since $\text{Opt} \setminus \partial\Lambda'_{+,e} \neq \emptyset$ (by assumption), from the containment $\text{Opt} \subset \Lambda'_{+,e}$, and from the convexity of the sets Opt and $\Lambda'_{+,e}$, follows by standard arguments that

$$\text{relint}(\text{Opt}) \cap \partial\Lambda'_{+,e} = \emptyset .$$

Thus, fixing optimal x not contained in $\text{relint}(\text{Opt})$, it remains to show $x \in \partial\Lambda'_{+,e}$, that is, to show $p'_e(x) = 0$.

Choose $y \in \text{relint}(\text{Opt})$, and let $y(t) := x + t(y - x)$ ($t \in \mathbb{R}$). Clearly, if $y(t) \in \text{Feas}$ then $y(t) \in \text{Opt}$. Thus, since x is in the relative boundary of Opt , it holds that $y(t) \notin \Lambda_+$ for $t < 0$.

Since the line segment with endpoints x and y is contained in $\partial\Lambda_+$, we have $p(y(t)) = 0$ for $0 \leq t \leq 1$. Since the only univariate polynomial with infinitely many roots is the polynomial that is identically zero, it follows $p(y(t)) = 0$ for *all* $t \in \mathbb{R}$. Hence, from the characterizations

$$\Lambda_{+,e}^{(k)} = \{z : p_e^{(j)}(z) \geq 0 \text{ for all } j = k, \dots, n-1\},$$

and the fact that $y(t) \notin \Lambda_+$ for $t < 0$, follows $p'_e(y(0)) = 0$, that is, $p'_e(x) = 0$, as desired. \square

For $e \in \Lambda_{++}$,

$$\text{define } q_e : \Lambda'_{++,e} \rightarrow \mathbb{R} \text{ by } q_e(x) := p(x)/p'_e(x).$$

As already mentioned, the purpose of this section is to develop a useful characterization of the set $\text{Opt} \setminus \partial\Lambda'_{+,e}$. The characterization is that the points in the set are precisely the optimal solutions to the following linearly constrained optimization problem:

$$\begin{aligned} \min_{x \in \Lambda'_{++,e}} \quad & -\ln c^*(e - x) - q_e(x) \\ \text{s.t.} \quad & Ax = b. \end{aligned} \tag{7.1}$$

Critical to achieving the characterization (and critical to characterization's relevance for computation) is Theorem 7, now restated and proved.

Theorem 7.2. *The rational function $q_e : \Lambda'_{++,e} \rightarrow \mathbb{R}$ is concave.*

Proof. Introduce a new variable t and let $P(x, t) := tp(x)$, a polynomial that is easily seen to be hyperbolic in direction $E = (e, 1)$. Let K' be the hyperbolicity cone for the derivative polynomial P'_E – thus, K' is the connected component of $S := \{(x, t) : p(x) + tp'_e(x) > 0\}$ containing E . We claim K' is precisely the interior of the epigraph for $-q_e : \Lambda'_{++,e} \rightarrow \mathbb{R}$. Since K' , being a hyperbolicity cone, is convex, establishing the claim will establish the theorem.

If $x \in \partial\Lambda'_{+,e}$ then $p'_e(x) = 0$ and, according to (3.5), $p(x) \leq 0$, so for no t do we have $(x, t) \in K'$. Thus, now assuming $x \in \Lambda'_{++,e}$ and $t > -q_e(x)$, to establish the claim it suffices to show there is a path in S from (x, t) to $E = (e, 1)$ (because that will imply (x, t) and E are in the same connected component of S).

Choose \bar{t} satisfying $\bar{t} > \max\{-q_e(sx + (1-s)e) : 0 \leq s \leq 1\}$ – the maximum exists because the line segment connecting x to e is contained in the convex set $\Lambda'_{++,e}$ and because p'_e is positive everywhere in $\Lambda'_{++,e}$. It is easily verified that a path in S from (x, t) to $(e, 1)$ is obtained with three line segments; the line segment between (x, t) and (x, \bar{t}) , the line segment between (x, \bar{t}) and (e, \bar{t}) , and the line segment between (e, \bar{t}) and $(e, 1)$. \square

We now restate and prove Theorem 8, the main result of this section.

Theorem 7.3. *If $e \in \Lambda_{++}$ then*

$$\text{Opt} \setminus \partial\Lambda'_{+,e} = \{x : x \text{ is optimal for (7.1)}\}$$

(possibly the empty set).

Proof. For x in $\Lambda'_{++,e}$ we have $Dp(x) \neq 0$ (because $0 < p'_e(x) = Dp(x)[e]$). Thus, from the characterization

$$(\partial\Lambda_+) \setminus \partial\Lambda'_{+,e} = \{x : p(x) = 0 \text{ and } p_e^{(j)}(x) > 0 \text{ for all } j = 1, \dots, n-1\}$$

(by (3.3)), it follows that for the convex optimization problem HP,

$$\begin{aligned} x &\in \text{Opt} \setminus \partial\Lambda'_{+,e} \\ &\Leftrightarrow \\ \lambda c^* &= Dp(x) + y^* A \quad \text{for some } \lambda > 0 \text{ and } y^* \\ Ax &= b \\ x &\in (\partial\Lambda_+) \setminus \partial\Lambda'_{+,e} . \end{aligned}$$

Observe that these conditions and homogeneity of p give

$$\lambda c^*(e - x) = Dp(x)[e - x] = p'_e(x) - n p(x) = p'_e(x) ,$$

that is,

$$\lambda = \frac{p'_e(x)}{c^*(e - x)} .$$

Clearly, then,

$$\left. \begin{aligned} x &\in \text{Opt} \setminus \partial\Lambda'_{+,e} \\ &\Leftrightarrow \\ \frac{p'_e(x)}{c^*(e-x)} c^* &= Dp(x) + y^* A \quad \text{for some } y^* \\ Ax &= b \\ x &\in (\partial\Lambda_+) \setminus \partial\Lambda'_{+,e} \end{aligned} \right\} \quad (7.2)$$

On the other hand, necessary and sufficient conditions for x to solve the convex optimization problem (7.1) are

$$\left. \begin{aligned} \frac{1}{c^*(e-x)} c^* - Dq_e(x) &= w^* A \quad \text{for some } w^* \\ Ax &= b \\ x &\in \Lambda'_{++,e} \end{aligned} \right\} \quad (7.3)$$

Observe that these conditions along with homogeneity of p and p'_e give

$$\begin{aligned} 1 &= \frac{1}{c^*(e-x)} c^*(e - x) \\ &= Dq_e(x)[e - x] \\ &= \frac{1}{p'_e(x)} \left(Dp(x)[e - x] - \frac{p(x)}{p'_e(x)} Dp'_e(x)[e - x] \right) \\ &= \frac{1}{p'_e(x)} \left(p'_e(x) - n p(x) - \frac{p(x)}{p'_e(x)} (p''_e(x) - (n-1)p'_e(x)) \right) \\ &= 1 - \frac{p(x)}{p'_e(x)} \left(1 + \frac{p''_e(x)}{p'_e(x)} \right) . \end{aligned}$$

Consequently, as $x \in \Lambda'_{++,e}$ implies $p'_e(x) > 0$ and $p''_e(x) > 0$, it must be for optimal x that $p(x) = 0$. Hence, in the necessary and sufficient optimality conditions (7.3), the containment $x \in \Lambda'_{++,e}$ can be replaced by $x \in (\partial\Lambda_+) \setminus \partial\Lambda'_{+,e}$. Since for x satisfying $p(x) = 0$ we also have $Dq_e(x) = \frac{1}{p'_e(x)} Dp(x)$, we see that the conditions (7.3) are equivalent to

$$\left. \begin{aligned} \frac{1}{c^*(e-x)} c^* - \frac{1}{p'_e(x)} Dp(x) &= w^* A \quad \text{for some } w^* \\ Ax &= b \\ x &\in (\partial\Lambda_+) \setminus \partial\Lambda'_{+,e} \end{aligned} \right\} \quad (7.4)$$

The optimality conditions (7.4) and (7.2) are identical. \square

We close this section with a restatement of Corollary 9.

Corollary 7.4. *If $1 \leq i \leq n-2$, $e \in \text{Swath}(i) \setminus \text{Core}(i)$ and*

$$f(x) := -\ln c^*(e - x) - \frac{p_e^{(i)}(x)}{p_e^{(i+1)}(x)} ,$$

then $x_e^{(i)}$ is the unique optimal solution for the convex optimization problem

$$\begin{array}{ll} \min_x & f(x) \\ \text{s.t.} & Ax = b . \end{array}$$

Moreover, $D^2f(x_e^{(i)})[v, v] > 0$ if v is not a scalar multiple of $x_e^{(i)}$ (in particular, if $v \neq 0$ satisfies $Av = 0$).

Proof. The first statement is immediate from Theorem 7.3. The statement regarding D^2f is immediate from Theorem 5.3(C). \square

As was noted in §2, the corollary's statement regarding D^2f implies quadratic convergence of Newton's method when initiated near to $x_e^{(i)}$. This is not pursued in the present paper.

8. Proof of Part I of Main Theorem

Recall that for the differential equation

$$\frac{d}{dt}e(t) = x_{e(t)}^{(i)} - e(t) , \quad e(0) = e \in \text{Swath}(i) \setminus \text{Core}(i) ,$$

we let $T = T(e(0))$ denote the time (possibly $T = \infty$) at which the trajectory terminates due either to reaching the boundary of $\text{Swath}(i) \setminus \text{Core}(i)$ or escaping to infinity. We say that $t \mapsto e(t)$ ($0 \leq t < T$) is a "maximal trajectory."

The primary goal of this section is to prove that the path $t \mapsto e(t)$, or the trajectory $t \mapsto x_{e(t)}^{(i)}$ (perhaps both), converges to optimality as $t \rightarrow T$.

As we shall see, the proof is reasonably straightforward when $T = \infty$, in which case the bounded trajectory $t \mapsto e(t)$ has all limit points in Opt .

Perhaps deserving of mention is that for all $e \in \text{Swath}(n-2) \setminus \text{Core}(n-2)$, it holds $T(e) = \infty$, and so our reasonably straightforward proof applies.

Unfortunately, the case $T < \infty$ is an entirely different matter. Here we show $t \mapsto e(t)$ converges to a unique point in $\text{Core}(i)$, and we show the path $t \mapsto x_{e(t)}^{(i)}$ is bounded, with all limit points in Opt . This proof is subtle and long. It would be great if a more direct proof were discovered.

Before restating the theorem and beginning the proof, we introduce a new use of differentials that plays a significant role in this section and later.

As was recalled for the reader in §3, the j^{th} differential at x for an analytic function f is the symmetric multilinear form defined on j -tuples of vectors $[v_1, \dots, v_j]$ by

$$D^j f(x)[v_1, \dots, v_j] := \frac{d}{dt_1} \cdots \frac{d}{dt_j} f(x + t_1 v_1 + \cdots + t_j v_j) |_{t_j=0} \cdots |_{t_1=0} .$$

Until now, significant roles have been played only by the differentials $D^j p_e^{(k)}(x)$ of the hyperbolic polynomials $x \mapsto p_e^{(k)}(x)$. Here, e is viewed as fixed, and x as the variable.

In this section, we sometimes need to view e as a variable, not just x . We need to differentiate once with respect to e , and j times with respect to x . For this we introduce the notation $D_e(D^j p_e^{(k)}(x))$ to represent the form that assigns to pairs of tuples $[u]$ and $[v_1, \dots, v_j]$ the value

$$\frac{d}{ds}(D^j p_{e+su}^{(i)}(x)[v_1, \dots, v_j])|_{s=0} .$$

(Here we differentiated at e in direction u only after differentiating at x in directions v_1, \dots, v_j , but the order of differentiation is immaterial, because $(e, x) \mapsto p_e^{(i)}(x)$ is a polynomial (in particular, is analytic)). Thus, there is good reason to also denote the form by, say, $D^j(D_e p_e^{(k)}(x))$.

When acting on a specific pair of tuples $[u]$ and $[v_1, \dots, v_j]$, we denote the assigned value by $D_e(D^j p_e^{(k)}(x)[v_1, \dots, v_j])[u]$ (or by $D^j(D_e p_e^{(k)}(x)[u])[v_1, \dots, v_j]$). The form is multilinear in u, v_1, \dots, v_j , and is symmetric in v_1, \dots, v_j .

The subscript on “ D_e ” does not refer to a specific hyperbolicity direction, but rather, to the derivative being taken with respect to the hyperbolicity direction. Thus, for example, $D_e(D^j p_{e(t)}^{(k)}(x))$ is the form that assigns to pairs $[u]$ and $[v_1, \dots, v_j]$ the value $\frac{d}{ds}(D^j p_{e(t)+su}^{(k)}(x)[v_1, \dots, v_j])|_{s=0}$.

We have occasion to fix the direction u , in which case results a form on tuples $[v_1, \dots, v_j]$; specifically,

$$[v_1, \dots, v_j] \mapsto D_e(D^j p_e^{(k)}(x)[v_1, \dots, v_j])[u],$$

a form we denote by $D_e(D^j p_e^{(i)}(x))[u]$. Similarly, if v_1, \dots, v_i and u are fixed, we use $D_e(D^j p_e^{(k)}(x)[v_1, \dots, v_i])[u]$ to denote the form

$$[w_1, \dots, w_{j-i}] \mapsto D_e(D^j p_e^{(k)}(x)[v_1, \dots, v_i, w_1, \dots, w_{j-i}])[u].$$

Time and again we rely on a relation between the forms $D_e(D^j p_e^{(k)}(x))$ and $D^{j+1} p_e^{(k-1)}(x)$ when $k \geq 1$:

$$\begin{aligned} D_e(D^j p_e^{(k)}(x)) &= D_e(D^{j+k} p(x) \underbrace{[e, \dots, e]}_{k \text{ times}}) \\ &= k D^{j+k} p(x) \underbrace{[e, \dots, e]}_{k-1 \text{ times}} \quad (\text{by multilinearity and symmetry}) \\ &= k D^{j+1} p_e^{(k-1)}(x). \end{aligned} \tag{8.1}$$

That is, for any \bar{e} and tuples $[u]$ and $[v_1, \dots, v_j]$,

$$D_e(D^j p_{\bar{e}}^{(k)}(x)[v_1, \dots, v_j])[u] = k D^{j+1} p_{\bar{e}}^{(k-1)}(x)[u, v_1, \dots, v_j]$$

(a consequence being that the form $D_e(D^j p_{\bar{e}}^{(k)}(x))$ is symmetric in u, v_1, \dots, v_j , not just in v_1, \dots, v_j).

Here is the restatement of the result to which this section is devoted, Part I of the Main Theorem:

Theorem 8.1. *Assume $1 \leq i \leq n-2$ and let $t \mapsto e(t)$ ($0 \leq t < T$) be a maximal trajectory generated by the dynamics $\dot{e}(t) = x_{e(t)}^{(i)} - e(t)$ beginning at $e(0) \in \text{Swath}(i) \setminus \text{Core}(i)$.*

- (A) *The trajectory $t \mapsto e(t)$ is bounded, and $t \mapsto c^* x_{e(t)}^{(i)}$ is strictly increasing, with val (the optimal value of HP) as the limit.*
- (B) *If $T = \infty$ then every limit point of the trajectory $t \mapsto e(t)$ lies in Opt .*
- (C) *If $T < \infty$ then the trajectory $t \mapsto e(t)$ has a unique limit point \bar{e} and $\bar{e} \in \text{Core}(i)$; moreover, the path $t \mapsto x_{e(t)}^{(i)}$ is bounded and each of its limit points lies in Opt .*

8.1. The case $T = \infty$. As mentioned, our proof of Theorem 8.1 when $T = \infty$ is much easier than when $T < \infty$. In this subsection, we focus on the case $T = \infty$, although the following proposition is important also to the case $T < \infty$.

Proposition 8.2. *Under the hypotheses of Theorem 8.1, the trajectory $t \mapsto e(t)$ is bounded, and $t \mapsto c^* x_{e(t)}^{(i)}$ is strictly increasing.*

Proof. Since $e(0) \in \text{Swath}(i) \setminus \text{Core}(i)$, the set $\text{Opt}_{e(0)}^{(i)}$ is bounded (indeed, consists of the single point $x_{e(0)}^{(i)}$), and hence so are the level sets $\{x \in \text{Feas}_{e(0)}^{(i)} : c^*x \leq \alpha\}$ ($\alpha \in \mathbb{R}$). As $\text{Feas} \subseteq \text{Feas}_{e(0)}^{(i)}$, the level sets $\{x \in \text{Feas} : c^*x \leq \alpha\}$ are bounded, too. Hence, since the trajectory $t \mapsto e(t)$ remains in Feas , and since (clearly) $t \mapsto c^*e(t)$ is decreasing (strictly monotonically), the trajectory lies entirely within a bounded region, that is, the trajectory is bounded.

It remains to prove that $t \mapsto c^*x_{e(t)}^{(i)}$ is strictly increasing. To ease notation, let $x(t) := x_{e(t)}^{(i)}$.

From $t \mapsto p_{e(t)}^{(i)}(x(t)) \equiv 0$ we find that

$$\begin{aligned} 0 &= \frac{d}{dt} p_{e(t)}^{(i)}(x(t)) \\ &= Dp_{e(t)}^{(i)}(x(t))[\dot{x}(t)] + D_e(p_{e(t)}^{(i)}(x(t)))[\dot{e}(t)] \\ &= Dp_{e(t)}^{(i)}(x(t))[\dot{x}(t)] + iDp_e^{(i-1)}(x(t))[\dot{e}(t)] \quad (\text{using (8.1)}) \\ &= Dp_{e(t)}^{(i)}(x(t))[\dot{x}(t)] + i \left((n-i+1)p_{e(t)}^{(i-1)}(x(t)) - p_{e(t)}^{(i)}(x(t)) \right), \end{aligned}$$

where the final equality is due to $\dot{e}(t) = x(t) - e(t)$ and, by (3.6),

$$Dp_{e(t)}^{(i-1)}(x(t))[\dot{x}(t)] = (n-i+1)p_{e(t)}^{(i-1)}(x(t)).$$

Thus, since $p_{e(t)}^{(i)}(x(t)) = 0$ and $p_{e(t)}^{(i-1)}(x(t)) < 0$ (by (3.4)), we have

$$Dp_{e(t)}^{(i)}(x(t))[\dot{x}(t)] > 0. \quad (8.2)$$

On the other hand, $x(t)$ satisfies the first-order condition

$$\lambda(t)c^* = Dp_{e(t)}^{(i)}(x(t)) + y(t)^*A \quad \text{for some } \lambda(t) > 0 \text{ and } y(t)^*.$$

Applying both sides to $\dot{x}(t)$, using $A\dot{x}(t) = 0$ and substituting (8.2), shows $c^*\dot{x}(t) > 0$. Thus, $t \mapsto c^*x(t)$ is strictly increasing. \square

We are now in position to easily prove Theorem 8.1 when $T = \infty$.

Theorem 8.3. *Under the hypotheses of Theorem 8.1, if $T = \infty$ then*

- $t \mapsto e(t)$ is bounded, with limit points lying in Opt , and
- $t \mapsto c^*x_{e(t)}^{(i)}$ increases strictly monotonically to val , the optimal value of HP.

Proof. Due to Proposition 8.2, it remains only to prove that both $t \mapsto c^*e(t)$ and $t \mapsto c^*x_{e(t)}^{(i)}$ converge to val as $t \rightarrow \infty$. However, the convergence of $c^*e(t)$ to val is immediate from

$$T = \infty, \quad c^*x_{e(t)}^{(i)} \leq \text{val} \leq c^*e(t) \quad \text{and} \quad \dot{e}(t) = x_{e(t)}^{(i)} - e(t). \quad (8.3)$$

Due to the monotonicity of $t \mapsto c^*x_{e(t)}^{(i)}$ (by Proposition 8.2), the convergence of $c^*x_{e(t)}^{(i)}$ to val also is immediate from (8.3) (indeed, otherwise at some time there would occur $c^*e(t) < \text{val}$, contradicting $e(t) \in \text{Feas}$). \square

8.2. The case $T < \infty$. For the challenging case $T < \infty$, we split most of the analysis into two propositions.

Proposition 8.4. *Under the hypotheses of Theorem 8.1, if $T < \infty$, then letting $x(t) := x_{e(t)}^{(i)}$, either*

- (A) *The trajectory $t \mapsto e(t)$ has a unique limit point, the limit point is contained in $\text{Core}(i)$, and the path $t \mapsto x(t)$ is bounded, with all limit points contained in Opt .*

or

- (B) It holds that $\liminf_{t \rightarrow T} \|x(t)\| = \infty$. Moreover, there exists a constant $\alpha < 0$ such that for all $0 \leq t < T$,

$$p_{e(t)}^{(i-1)}(x(t)) \leq \alpha \|x(t)\|^{n-i+1}.$$

Proposition 8.5. *Under the hypotheses of Theorem 8.1, and regardless of whether T is finite, letting $x(t) := x_{e(t)}^{(i)}$, for all $0 \leq t < T$ we have*

$$\frac{d}{dt} \ln \frac{\left(-p_{e(t)}^{(i-1)}(x(t))\right)^{n-i}}{\left(p_{e(t)}^{(i+1)}(x(t))\right)^{n-i+1}} \leq C_1 \frac{p_{e(t)}^{(i-1)}(x(t))}{p_{e(t)}^{(i+1)}(x(t))} + C_2 \frac{p_{e(t)}^{(i-2)}(x(t))}{p_{e(t)}^{(i-1)}(x(t))} + C_3,$$

where

$$C_1 = i(n-i+1)^2, \quad C_2 = (i-1)(n-i)(n-i+2) \quad \text{and} \quad C_3 = 2n-i+1$$

(and where for the case $i = 1$ we define $p_{e(t)}^{(-1)} \equiv 0$).

Before proving the two propositions, we show how they complete the proof of Theorem 8.1.

Theorem 8.6. *Under the hypotheses of Theorem 8.1, if $T < \infty$ then*

- $t \mapsto e(t)$ has a unique limit point, the limit point is contained in $\text{Core}(i)$, and
- $t \mapsto x_{e(t)}^{(i)}$ is bounded, having all limit points contained in Opt ;
moreover, $t \mapsto c^* x_{e(t)}^{(i)}$ increases strictly monotonically (to val, the optimal value of HP).

Proof. We already know $t \mapsto c^* x_{e(t)}^{(i)}$ is strictly increasing. Completing the proof thus amounts to showing case (A) of Proposition 8.4 holds. Hence, it suffices to assume that case (B) of Proposition 8.4 holds, and then show Proposition 8.5 yields a contradiction. Thus, assume case (B) holds.

Since the trajectory $t \mapsto e(t)$ is bounded (by Proposition 8.2), we can choose a compact set K containing the entire trajectory (we assume nothing about K other than it is compact). For $j = i-2, i-1, i+1$, define

$$\gamma_j := \max\{|p_{\bar{e}}^{(j)}(x)| : \bar{e} \in K \text{ and } \|x\| = 1\}$$

(let $\gamma_{i-2} := 0$ when $i = 1$). Since $p_{\bar{e}}^{(j)}$ is homogeneous of degree $n-j$, we have

$$|p_{\bar{e}}^{(j)}(x)| \leq \gamma_j \|x\|^{n-j} \quad \text{for all } \bar{e} \in K \text{ and all } x.$$

Let $x(t) := x_{e(t)}^{(i)}$.

Using the assumed bound $p_{e(t)}^{(i-1)}(x(t)) \leq \alpha \|x(t)\|^{n-i+1}$ – keep in mind especially that α is negative – and using the positivity of $p_{e(t)}^{(i+1)}(x(t))$ (according to (4.2)), from the inequality of Proposition 8.5 we find for all $0 \leq t < T$,

$$\frac{d}{dt} \ln \frac{\left(-p_{e(t)}^{(i-1)}(x(t))\right)^{n-i}}{\left(p_{e(t)}^{(i+1)}(x(t))\right)^{n-i+1}} \leq C_1 \frac{\alpha}{\gamma_{i+1}} \|x(t)\|^2 + C_2 \frac{\gamma_{i-2}}{|\alpha|} \|x(t)\| + C_3.$$

Hence, there exists a value r for which

$$\|x(t)\| > r \quad \Rightarrow \quad \frac{d}{dt} \ln \frac{\left(-p_{e(t)}^{(i-1)}(x(t))\right)^{n-i}}{\left(p_{e(t)}^{(i+1)}(x(t))\right)^{n-i+1}} < 0.$$

Since $\liminf_{t \rightarrow T} \|x(t)\| = \infty$ (by assumption), it follows that

$$\sup_{0 \leq t < T} \frac{\left(-p_{e(t)}^{(i-1)}(x(t))\right)^{n-i}}{\left(p_{e(t)}^{(i+1)}(x(t))\right)^{n-i+1}} < \infty.$$

However,

$$\begin{aligned} \frac{\left(-p_{e(t)}^{(i-1)}(x(t))\right)^{n-i}}{\left(p_{e(t)}^{(i+1)}(x(t))\right)^{n-i+1}} &\geq \frac{\left(|\alpha| \|x(t)\|^{n-i+1}\right)^{n-i}}{(\gamma_{i+1} \|x(t)\|^{n-i-1})^{n-i+1}} \\ &= \frac{|\alpha|^{n-i}}{\gamma_{i+1}^{n-i-1}} \|x(t)\|^{n-i+1} \rightarrow \infty, \end{aligned}$$

a contradiction, thus concluding the proof of the theorem (except for proving the two propositions). \square

8.2.1. Proving the first of the two propositions. Now we begin proving Proposition 8.4. The proof relies on three lemmas.

Lemma 8.7. *Assume Opt is a bounded set and $1 \leq i \leq n-2$. Let $\{e_j\}$ be a sequence of derivative directions converging to $e \in \Lambda_{++}$, and assume $x_j \in \text{Opt}_{e_j}^{(i)}$.*

- (A) *If $x_j \rightarrow x$, then $x \in \text{Opt}_e^{(i)}$.*
- (B) *If $\{x_j\}$ is an unbounded set, then $\text{Opt}_e^{(i)} = \emptyset$ and $\liminf \|x_j\| = \infty$.*
- (C) *If $\{x_j\}$ is an unbounded set and $\limsup c^* x_j > -\infty$, then $x_j/\|x_j\|$ has exactly one limit point d ; moreover, d satisfies $p_e^{(i-1)}(d) < 0$.*

Proof. First assume $x_j \rightarrow x$. Then $p_e^{(k)}(x) = \lim p_{e_j}^{(k)}(x_j)$ for all k . However, by (3.1), $p_{e_j}^{(k)}(x_j) \geq 0$ for all $k = i, \dots, n-1$. Thus, $p_e^{(k)}(x) \geq 0$ for all $k = i, \dots, n-1$. Hence, again invoking (3.1), $x \in \Lambda_{+,e}^{(i)}$. Since, trivially, $Ax = b$, we thus see that $x \in \text{Feas}_e^{(i)}$.

To prove that x not only is feasible but is optimal, we assume otherwise and obtain a contradiction. Thus, assume $\bar{x} \in \text{Feas}_e^{(i)}$ satisfies $c^* \bar{x} < c^* x$. By nudging \bar{x} towards the strictly feasible point e , we may assume \bar{x} is strictly feasible in addition to satisfying $c^* \bar{x} < c^* x$. However, by (3.2), strict feasibility of \bar{x} implies $p_e^{(k)}(\bar{x}) > 0$ for all $k = i, \dots, n-1$. Since $p_{e_j}^{(k)}(\bar{x}) \rightarrow p_e^{(k)}(\bar{x})$, we thus have, again using (3.1), that $\bar{x} \in \text{Feas}_{e_j}^{(i)}$ for $j > J$ (some J). But then

$$c^* x > c^* \bar{x} \geq \lim c^* x_j = c^* x,$$

a contradiction. Hence, x is optimal, and assertion (A) of the lemma is established.

Now assume $\{x_j\}$ is unbounded. Choose a subsequence $\{x_{j_\ell}\}$ for which $\liminf_{\ell \rightarrow \infty} \|x_{j_\ell}\| = \infty$ and let d be a limit point of $\{x_{j_\ell}/\|x_{j_\ell}\|\}$. Then d is a feasible direction for $\text{HP}_e^{(i)}$, that is, satisfies $Ad = 0$ and $d \in \Lambda_{+,e}^{(i)}$. (That $Ad = 0$ is trivial. That $d \in \Lambda_{+,e}^{(i)}$ follows from $p_e^{(k)}(d) = \lim_{\ell \rightarrow \infty} p_{e_{j_\ell}}^{(k)}(x_{j_\ell}/\|x_{j_\ell}\|) \geq 0$ for $k = i, \dots, n-1$, the inequality being due to $x_{j_\ell} \in \Lambda_{+,e_{j_\ell}}^{(i)}$ along with homogeneity of $p_{e_{j_\ell}}^{(i)}$.) The feasible direction d satisfies $c^* d \leq 0$ (because $e \in \text{Feas}_{e_{j_\ell}}^{(i)}$ – and hence $\limsup_{\ell \rightarrow \infty} c^* x_{j_\ell} < \infty$ – and because $\liminf_{\ell \rightarrow \infty} \|x_{j_\ell}\| = \infty$).

Clearly, now, if $\text{Opt}_e^{(i)} \neq \emptyset$ then $\text{Opt}_e^{(i)}$ is unbounded (in direction d), implying by Corollary 4.2(B) that $\text{Opt}_e^{(i)} = \text{Opt}$. But this would contradict our assumption that Opt is bounded. Thus, $\text{Opt}_e^{(i)} = \emptyset$.

To conclude the proof of assertion (B) of the lemma, it remains to show $\liminf \|x_j\| = \infty$. But if this identity did not hold then $\{x_j\}$ would have a limit point, and thus,

by assertion (A) of the lemma, we would have $\text{Opt}_e^{(i)} \neq \emptyset$, contradicting what we just proved. The proof of assertion (B) of the lemma is now complete.

Now we prove assertion (C). Assume $\{x_j\}$ is unbounded and $\limsup c^*x_j > -\infty$.

With $\{x_j\}$ unbounded, we already know from above that every limit point d of $\{x_j/\|x_j\|\}$ is a feasible direction for $\text{HP}_e^{(i)}$ (that is, satisfies $Ad = 0$ and $d \in \Lambda_{+,e}^{(i)}$) and $c^*d \leq 0$.

We claim, however, that every feasible direction d for $\text{HP}_e^{(i)}$ satisfies $c^*d \geq 0$. Indeed, otherwise the optimal objective value of $\text{HP}_e^{(i)}$ would be unbounded and hence there would exist $\bar{x} \in \text{relint}(\text{Feas}_e^{(i)})$ satisfying $c^*\bar{x} < \limsup c^*x_j$ (using the assumption $\limsup c^*x_j > -\infty$). But $\bar{x} \in \text{Feas}_{e_j}^{(i)}$ for sufficiently large j (because $0 < p_e^{(k)}(\bar{x}) = \lim p_{e_j}^{(k)}(\bar{x})$ for $k = i, \dots, n-1$), contradicting that x_j is optimal for $\text{HP}_{e_j}^{(i)}$ for all j .

Clearly, from the two preceding paragraphs, every limit point of $\{x_j/\|x_j\|\}$ is contained in $D := \{d \in \Lambda_{+,e}^{(i)} : Ad = 0 \text{ and } c^*d = 0\}$. To show $\{x_j/\|x_j\|\}$ has a unique limit point it thus suffices to show D consists of a single ray.

We claim $D \cap \Lambda_{+,e}^{(i)} = \emptyset$. Otherwise there would exist $d \in D$ satisfying $p_e^{(k)}(d) > 0$ for $k = i, \dots, n-1$, which from continuity would imply d to be a feasible direction for $\text{HP}_{e_j}^{(i)}$ for all $j > J$ (some J). But then $\text{Opt}_{e_j}^{(i)}$ would be an unbounded set for $j > J$, implying by Corollary 4.2(B) that $\text{Opt}_{e_j}^{(i)} = \text{Opt}$, and thus contradicting the assumed boundedness of Opt . Hence, it is indeed the case that $D \cap \Lambda_{+,e}^{(i)} = \emptyset$.

It is clear now that $D = \{d \in \partial\Lambda_{+,e}^{(i)} : Ad = 0 \text{ and } c^*d = 0\}$. Thus, D is contained in a boundary face of $\Lambda_{+,e}^{(i)}$. Hence, D consists of a single ray or is a subset of Λ_+ , by Corollary 4.2(B). But if D were a subset of Λ_+ then each $d \in D$ would be a feasible direction for HP , which, with $c^*d = 0$, would imply Opt to be unbounded, a contradiction to the assumed boundedness of Opt . Thus, D consists of a single ray, and so $\{x_j/\|x_j\|\}$ has a unique limit point d (and d satisfies $d \in \partial\Lambda_{+,e}^{(i)}$, $Ad = 0$, $c^*d = 0$).

Lastly, we prove $p_e^{(i-1)}(d) < 0$.

As $x_j \in \partial\Lambda_{+,e_j}^{(i)}$, we have $p_{e_j}^{(i-1)}(x_j) \leq 0$ (by (3.5)). Thus, by homogeneity and continuity, $p_e^{(i-1)}(d) \leq 0$. However, since $d \in \partial\Lambda_{+,e}^{(i)}$, if d satisfied $p_e^{(i-1)}(d) = 0$, then we would have $d \in \Lambda_{+,e}^{(i-1)}$, and hence $d \in \Lambda_+$ (by Theorem 4.1(A)) – so d would be a feasible direction for HP , and hence Opt would be unbounded (using $c^*d = 0$), a contradiction. Thus, $p_e^{(i-1)}(d) < 0$, completing the proof of the lemma. \square

Lemma 8.8. *Assume Opt is a bounded set, $e \in \Lambda_{++}$ and $1 \leq i \leq n-2$. Assume there exist sequences $e_j \in \text{Swath}(i)$ and $x_j \in \text{Opt}_{e_j}^{(i)}$ for which $e_j \rightarrow e$, $\{x_j\}$ is unbounded and $\limsup c^*x_j > -\infty$.*

(A) *There exists a constant $\alpha < 0$ and an open neighborhood U of e such that for all $\bar{e} \in U \cap \text{Swath}(i)$ and $\bar{x} \in \text{Opt}_{\bar{e}}^{(i)}$,*

$$p_{\bar{e}}^{(i-1)}(\bar{x}) < \alpha \|\bar{x}\|^{n-i+1}.$$

(B) *For each $\ell \in \mathbb{R}$, there exists an open neighborhood $V(\ell)$ of e such that for all $\bar{e} \in V(\ell) \cap \text{Swath}(i)$ and $\bar{x} \in \text{Opt}_{\bar{e}}^{(i)}$, it holds that $\|\bar{x}\| \geq \ell$.*

Proof. Parts (A) and (B) of Lemma 8.7 together imply that if for one sequence $\{e_j\} \subset \text{Swath}(i)$ there exist $x_j \in \text{Opt}_{e_j}^{(i)}$ for which $\{x_j\}$ is unbounded – where $e_j \rightarrow e$ – then for every sequence $\{e_j\} \subset \text{Swath}(i)$ converging to e , and every choice of $x_j \in \text{Opt}_{e_j}^{(i)}$, we have $\liminf \|x_j\| = \infty$. Part (B) of the present lemma easily follows.

Towards proving part (A), let e_j and x_j be sequences as in the statement of the present lemma. Part (C) of Lemma 8.7 then shows $x_j/\|x_j\|$ has a unique limit point d .

We claim $\bar{x}_j/\|\bar{x}_j\| \rightarrow d$ for any sequences \bar{e}_j and $\bar{x}_j \in \text{Opt}_{\bar{e}_j}^{(i)}$ satisfying $\bar{e}_j \rightarrow e$. Indeed, the intermingled sequences $e_1, \bar{e}_1, e_2, \bar{e}_2, \dots$ and $x_1, \bar{x}_1, x_2, \bar{x}_2, \dots$ clearly satisfy the hypotheses of part (C) of Lemma 8.7, and hence the normalized sequence

$$x_1/\|x_1\|, \bar{x}_1/\|\bar{x}_1\|, x_2/\|x_2\|, \bar{x}_2/\|\bar{x}_2\|, \dots$$

has a unique limit point, which, of course, is d , thus establishing the claim.

Choosing, say, $\alpha = \frac{1}{2} p_e^{(i-1)}(d)$, part (A) of the present lemma is now a consequence of the continuity of the polynomial $(\bar{e}, \bar{x}) \mapsto p_{\bar{e}}^{(i-1)}(\bar{x})$, and the fact that $x \mapsto p_{\bar{e}}^{(i-1)}(x)$ is homogeneous of degree $n - i + 1$ for all \bar{e} . \square

Throughout the remainder of the section,

- assume $1 \leq i \leq n - 2$,
- assume $t \mapsto e(t)$ ($0 \leq t < T$) is a maximal trajectory defined by the differential equation $\dot{e}(t) = x_{e(t)}^{(i)} - e(t)$, starting at $e(0) \in \text{Swath}(i) \setminus \text{Core}(i)$,
- assume $T < \infty$, and
- to ease notation, let $x(t) := x_{e(t)}^{(i)}$.

The third (and final) lemma for proving Proposition 8.4 shows that when the trajectory $t \mapsto e(t)$ terminates at finite time T , the termination is not due to having reached the boundary of Feas.

Lemma 8.9. $\{e(t) : 0 \leq t < T\} \subset \Lambda_{++}$

Proof. It suffices to show $p(e(t)) \geq p(e(0))e^{-nt}$ for all t , and hence suffices to show $\frac{d}{dt}p(e(t)) \geq -np(e(t))$, that is, suffices to show $Dp(e(t))[\dot{e}(t)] \geq -np(e(t))$. Since $\dot{e}(t) = x(t) - e(t)$ and since $Dp(e(t))[e(t)] = np(e(t))$ (by (3.6)), it suffices to show $Dp(e(t))[x(t)] \geq 0$.

However, by (3.7), for all \bar{e} , $Dp(\bar{e})$ is a positive multiple of $Dp_{\bar{e}}^{(n-1)}(\bar{e})$. Since for all x , $Dp_{\bar{e}}^{(n-1)}(\bar{e})[x] = p_{\bar{e}}^{(n-1)}(x)$ (because $x \mapsto p_{\bar{e}}^{(n-1)}(x)$ is linear), we have $Dp(\bar{e})[x] \geq 0$ if and only if x is in the cone $\Lambda_{+, \bar{e}}^{(n-1)}$. But $x(t) \in \Lambda_{+, e(t)}^{(i)} \subseteq \Lambda_{+, e(t)}^{(n-1)}$, thus concluding the proof. \square

Proof of Proposition 8.4. To rely on Lemmas 8.7 and 8.8, we need Opt to be a bounded set. Boundedness of Opt , however, was established in Proposition 8.2 as a simple consequence of $\text{Swath}(i) \setminus \text{Core}(i)$ being nonempty.

Let \mathcal{L} be the set of limit points for the trajectory $t \mapsto e(t)$ as $t \rightarrow T$. By Lemma 8.9, $\mathcal{L} \subset \Lambda_{++}$. Thus, if $e \in \mathcal{L}$, then either $e \in \text{Core}(i)$ (that is, $\text{Opt}_e^{(i)} = \text{Opt}$) or $e \in \Lambda_{++} \setminus \text{Swath}(i)$ (in which case $\text{Opt}_e^{(i)} = \emptyset$).

Assume first that \mathcal{L} consists of a single point e .

Then, by Lemma 8.7(A), each limit point of the path $t \mapsto x(t)$ lies in $\text{Opt}_e^{(i)}$. If the path is bounded, the set of limit points is nonempty, hence $\text{Opt}_e^{(i)} \neq \emptyset$, and thus, from above, $e \in \text{Core}(i)$. Clearly, then, if the path $t \mapsto x(t)$ is bounded, case (A) of the proposition holds.

On the other hand, if the path $t \mapsto x(t)$ is unbounded, then Lemma 8.8 implies $\liminf_{t \rightarrow T} \|x(t)\| = \infty$, and implies there exists $0 \leq \bar{t} < T$ and $\alpha < 0$ satisfying

$$p_{e(t)}^{(i-1)}(x(t)) \leq \alpha \|x(t)\|^{n-i+1} \quad \text{for all } \bar{t} \leq t < T \quad (8.4)$$

(here we make use of the fact that $t \mapsto c^*x(t)$ is strictly increasing (Proposition 8.2), so that the hypothesis $\limsup c^*x_j > -\infty$ of the lemma is clearly fulfilled). Hence, from the compactness of the closed interval $[0, \bar{t}]$ and the fact that for all t , both

$x(t) \neq 0$ (because $b \neq 0$) and $p_{e(t)}^{(i-1)}(x(t)) < 0$ (by (3.4)), a perhaps larger (but still negative) value of α satisfies

$$p_{e(t)}^{(i-1)}(x(t)) \leq \alpha \|x(t)\|^{n-i+1} \quad \text{for all } 0 \leq t < T. \quad (8.5)$$

Hence, in all, if the path $t \mapsto x(t)$ is unbounded, then case (B) of the proposition holds, concluding consideration of the case that \mathcal{L} consists of a single point.

For the remainder of the proof assume \mathcal{L} contains more than one point. We show case (B) of the proposition holds. We claim that for this it suffices to show \mathcal{L} is compact, and that each $e \in \mathcal{L}$ satisfies the hypotheses of Lemma 8.8. Indeed, compactness and Lemma 8.8(A) imply \mathcal{L} can be covered by finitely many open sets U_j for which there exist $\alpha_j < 0$ with the property that for all $\bar{e} \in U_j \cap \text{Swath}(i)$ and $\bar{x} \in \text{Opt}_{\bar{e}}^{(i)}$, it holds $p_{\bar{e}}^{(i-1)}(\bar{x}) < \alpha_j \|\bar{x}\|^{n-i+1}$. Consequently, letting $\alpha := \max_j \alpha_j$, there exists $0 \leq \bar{t} < T$ for which (8.4) holds. Then, as in the preceding paragraph, for a possibly larger (but still negative) value of α , (8.5) holds. To conclude proving the claim, it remains only to show $\liminf_{t \rightarrow T} \|x(t)\| = \infty$. This, however, is easily accomplished by covering \mathcal{L} , for each $\ell \in \mathbb{R}$, by finitely many open sets $V(\ell)_j$ as appear in Lemma 8.8(B).

Thus, to complete the proof of Proposition 8.4, it remains only to show \mathcal{L} is compact, and that each $e \in \mathcal{L}$ satisfies the hypotheses of Lemma 8.8.

That \mathcal{L} is compact is trivial – indeed, the trajectory $t \mapsto e(t)$ is bounded, by Proposition 8.2.

Since \mathcal{L} consists of more than one point, for each $e \in \mathcal{L}$ and each open neighborhood W of e , the Euclidean arc length of $\{e(t) : 0 \leq t < T\} \cap W$ is infinite. Thus, since $T < \infty$, for each $e \in \mathcal{L}$ there exists an sequence $t_1 < t_2 < \dots$ satisfying $e(t_j) \rightarrow e$ and $\|\dot{e}(t_j)\| \rightarrow \infty$. But $\dot{e}(t_j) = x(t_j) - e(t_j)$, so $\|x(t_j)\| \rightarrow \infty$. Since, additionally, $\limsup c^*x(t_j) > -\infty$ (because $t \mapsto c^*x(t)$ is strictly increasing), e thus clearly satisfies the hypotheses of Lemma 8.8. \square

8.2.2. Proving the second of the two propositions. In proving Proposition 8.5, continue to

- assume $1 \leq i \leq n-2$,
- assume $t \mapsto e(t)$ ($0 \leq t < T$) is a maximal trajectory defined by the differential equation $\dot{e}(t) = x_{e(t)}^{(i)} - e(t)$, starting at $e(0) \in \text{Swath}(i) \setminus \text{Core}(i)$,
- assume $T < \infty$, and
- to ease notation, let $x(t) := x_{e(t)}^{(i)}$.

Also, for every e , define $q_e := p_e^{(i)} / p_e^{(i+1)}$.

The proof of Proposition 8.5 depends on two lemmas. For understanding the first lemma, recall the identity given as (8.1), that is,

$$D_e(D^j p_e^{(k)}(x)) = k D^{j+1} p_e^{(k-1)}(x). \quad (8.6)$$

Lemma 8.10. *For any e and x satisfying $p_e^{(i)}(x) = 0 \neq p_e^{(i+1)}(x)$, we have*

$$\begin{aligned} D_e(Dq_e(x))[x - e] &= Dq_e(x) \\ &+ \frac{i(n-i)}{p_e^{(i+1)}(x)} Dp_e^{(i-1)}(x) \\ &- \frac{i(n-i+1)p_e^{(i-1)}(x)}{(p_e^{(i+1)}(x))^2} Dp_e^{(i+1)}(x). \end{aligned}$$

Proof. For an arbitrary vector Δe and for any x at which q_e is defined (i.e., any x satisfying $p_e^{(i+1)}(x) \neq 0$), use of (8.6) gives

$$\begin{aligned} D_e(Dq_e(x))[\Delta e] &= D_e \left(\frac{1}{p_e^{(i+1)}(x)} Dp_e^{(i)}(x) - \frac{p_e^{(i)}(x)}{p_e^{(i+1)}(x)^2} Dp_e^{(i+1)}(x) \right) [\Delta e] \\ &= \frac{i}{p_e^{(i+1)}(x)} D^2 p_e^{(i-1)}(x) [\Delta e] \\ &\quad - \frac{i+1}{(p_e^{(i+1)}(x))^2} (Dp_e^{(i)}(x) [\Delta e]) Dp_e^{(i)}(x) \\ &\quad - \frac{i}{(p_e^{(i+1)}(x))^2} (Dp_e^{(i-1)}(x) [\Delta e]) Dp_e^{(i+1)}(x) \\ &\quad - \frac{2(i+1)p_e^{(i)}(x)}{(p_e^{(i+1)}(x))^3} (Dp_e^{(i)}(x) [\Delta e]) Dp_e^{(i+1)}(x) \\ &\quad - \frac{(i+1)p_e^{(i)}(x)}{(p_e^{(i+1)}(x))^2} D^2 p_e^{(i)}(x) [\Delta e] . \end{aligned}$$

Substitute $\Delta e = x - e$, then use obvious identities to get rid of “[e]” (e.g., $Dp_e^{(i)}(x)[e] = p_e^{(i+1)}(x)$, $D^2 p_e^{(i-1)}(x)[e] = Dp_e^{(i)}(x)$), and use homogeneity to get rid of “[x]” – specifically, use (3.6). Finally, substitute $p_e^{(i)}(x) = 0$, and $\frac{1}{p_e^{(i+1)}(x)} Dp_e^{(i)}(x) = Dq_e(x)$ (because $p_e^{(i)}(x) = 0$), thereby concluding the proof. \square

Lemma 8.11. *For all $0 \leq t < T$, there exists $y(t)^*$ satisfying*

$$\begin{aligned} &D^2 q_{e(t)}(x(t))[\dot{x}(t)] - \frac{c^* \dot{x}(t)}{(c^*(e(t) - x(t)))^2} c^* \\ &= \frac{ip_{e(t)}^{(i-1)}(x(t))}{p_{e(t)}^{(i+1)}(x(t))} \left((n-i+1) \frac{Dp_{e(t)}^{(i+1)}(x(t))}{p_{e(t)}^{(i+1)}(x(t))} - (n-i) \frac{Dp_{e(t)}^{(i-1)}(x(t))}{p_{e(t)}^{(i-1)}(x(t))} \right) + y(t)^* A . \end{aligned}$$

Proof. Corollary 7.4 shows $x(t)$ is optimal for the convex optimization problem

$$\begin{aligned} \min_x \quad & -\ln c^*(e(t) - x) - q_{e(t)}(x) \\ \text{s.t.} \quad & Ax = b \end{aligned}$$

and hence there exists $w(t)^*$ satisfying the first-order condition

$$\frac{1}{c^*(e(t) - x(t))} c^* - Dq_{e(t)}(x(t)) = w(t)^* A . \quad (8.7)$$

Since $t \mapsto e(t)$ and $t \mapsto x_{e(t)}^{(i)}$ are analytic, so is $t \mapsto w^*(t)$ (using that A is surjective, by assumption). Differentiating in t gives

$$\frac{c^*(\dot{e}(t) - \dot{x}(t))}{(c^*(e(t) - x(t)))^2} c^* + D^2 q_{e(t)}(x(t))[\dot{x}(t)] + D_{e(t)}(Dq_{e(t)}(x(t)))[\dot{e}(t)] = -\dot{w}(t)^* A .$$

To complete the proof, substitute

$$\dot{e}(t) = x(t) - e(t) \quad \text{and} \quad \frac{1}{c^*(e(t) - x(t))} c^* = Dq_{e(t)}(x(t)) + w(t)^* A \quad (\text{by (8.7)}) ,$$

and then use Lemma 8.10 to substitute for $D_{e(t)}(Dq_{e(t)}(x(t)))[x(t) - e(t)]$. \square

Proof of Proposition 8.5. To temper notation, for arbitrary t satisfying $0 \leq t < T$, let

$$e := e(t), \quad x := x(t), \quad \dot{e} := \dot{e}(t) \quad \text{and} \quad \dot{x} := \dot{x}(t).$$

Observe that

$$\begin{aligned}
& \frac{d}{dt} \ln \frac{\left(-p_{e(t)}^{(i-1)}(x(t))\right)^{n-i}}{\left(p_{e(t)}^{(i+1)}(x(t))\right)^{n-i+1}} \\
&= (n-i) \frac{d}{dt} \ln \left(-p_{e(t)}^{(i-1)}(x(t))\right) - (n-i+1) \frac{d}{dt} \ln p_{e(t)}^{(i+1)}(x(t)) \\
&= \frac{n-i}{p_e^{(i-1)}(x)} \left(Dp_e^{(i-1)}(x)[\dot{x}] + D_e(p_e^{(i-1)}(x))[\dot{e}]\right) \\
&\quad - \frac{n-i+1}{p_e^{(i+1)}(x)} \left(Dp_e^{(i+1)}(x)[\dot{x}] + D_e(p_e^{(i+1)}(x))[\dot{e}]\right).
\end{aligned}$$

Using Lemma 8.11 to substitute for $\frac{n-i}{p_e^{(i-1)}(x)} Dp_e^{(i-1)}(x) - \frac{n-i+1}{p_e^{(i+1)}(x)} Dp_e^{(i+1)}(x)$ gives

$$\begin{aligned}
& \frac{d}{dt} \ln \frac{\left(-p_{e(t)}^{(i-1)}(x(t))\right)^{n-i}}{\left(p_{e(t)}^{(i+1)}(x(t))\right)^{n-i+1}} \\
&= -\frac{p_e^{(i+1)}(x)}{ip_e^{(i-1)}(x)} \left(D^2 q_e(x)[\dot{x}, \dot{x}] - \left(\frac{c^* \dot{x}}{c^*(e-x)}\right)^2\right) \\
&\quad + (n-i) \frac{D_e(p_e^{(i-1)}(x))[\dot{e}]}{p_e^{(i-1)}(x)} - (n-i+1) \frac{D_e(p_e^{(i+1)}(x))[\dot{e}]}{p_e^{(i+1)}(x)} \quad (\text{using } y^*(t)A\dot{x}(t) = 0) \\
&\leq \frac{p_e^{(i+1)}(x)}{ip_e^{(i-1)}(x)} \left(\frac{c^* \dot{x}}{c^*(e-x)}\right)^2 + (n-i) \frac{D_e(p_e^{(i-1)}(x))[\dot{e}]}{p_e^{(i-1)}(x)} - (n-i+1) \frac{D_e(p_e^{(i+1)}(x))[\dot{e}]}{p_e^{(i+1)}(x)},
\end{aligned}$$

where the inequality is due to the combination of $p_e^{(i+1)}(x)/p_e^{(i-1)}(x)$ being negative (according to (3.4) and (4.2)) and $D^2 q_e(x)$ being negative semidefinite (Theorem 7.2).

To complete the proof, simply substitute the following expressions which result from use of $\dot{e} = x - e$, (3.6) and (8.1):

$$\begin{aligned}
D_e(p_e^{(i-1)}(x))[\dot{e}] &= (i-1)Dp_e^{(i-2)}(x)[x-e] \\
&= (i-1) \left((n-i+2)p_e^{(i-2)}(x) - p_e^{(i-1)}(x)\right),
\end{aligned}$$

$$\begin{aligned}
D_e(p_e^{(i+1)}(x))[\dot{e}] &= (i+1)Dp_e^{(i)}(x)[x-e] \\
&= (i+1) \left(0 - p_e^{(i+1)}(x)\right)
\end{aligned}$$

– substitute, also,

$$\begin{aligned}
\frac{c^* \dot{x}}{c^*(e-x)} &= Dq_e(x)[\dot{x}] \quad (\text{first-order condition arising from Corollary 7.4}) \\
&= \frac{Dp_e^{(i)}(x)[\dot{x}]}{p_e^{(i+1)}(x)} \quad (\text{using } p_e^{(i)}(x) = 0) \\
&= -\frac{D_e(p_e^{(i)}(x))[\dot{e}]}{p_e^{(i+1)}(x)} \quad (\text{because } \frac{d}{dt} p_e^{(i)}(x_{e(t)}) = 0) \\
&= -i \frac{Dp_e^{(i-1)}(x)[x-e]}{p_e^{(i+1)}(x)} \\
&= -i \frac{(n-i+1)p_e^{(i-1)}(x) - 0}{p_e^{(i+1)}(x)}.
\end{aligned}$$

□

9. Proof of Part II of the Main Theorem

Recall the optimization problem dual to HP:

$$\left. \begin{array}{ll} \sup_{y^*, s^*} & y^* b \\ \text{s.t.} & y^* A + s^* = c^* \\ & s^* \in \Lambda_+^* \end{array} \right\} \text{HP}^*,$$

where Λ_+^* is the cone dual to Λ_+ . Recall that just as a matter of definitions, the optimal value of HP^* satisfies $\text{val}^* \leq \text{val}$ (“weak duality”), where val is the optimal value of HP.

Recall that a pair (y^*, s^*) satisfying the constraints is said to be “strictly” feasible if $s^* \in \text{int}(\Lambda_+^*)$ (interior).

For $e \in \text{Swath}(i) \setminus \text{Core}(i)$, define

$$s_e^{(i)} := \frac{c^*(e - x)}{p_e^{(i+1)}(x)} Dp_e^{(i)}(x) \quad \text{where } x = x_e^{(i)}.$$

We now restate and prove Part II of the Main Theorem.

Theorem 9.1. *Assume $1 \leq i \leq n - 2$ and let $\{e(t) : 0 \leq t < T\}$ be a maximal trajectory for the dynamics $\dot{e}(t) = x_{e(t)}^{(i)} - e(t)$ starting at $e(0) \in \text{Swath}(i) \setminus \text{Core}(i)$. Then $y^* A + s_{e(t)}^{(i)} = c^*$ has a unique solution $y^* = y_{e(t)}^{(i)}$, and the pair $(y_{e(t)}^{(i)}, s_{e(t)}^{(i)})$ is strictly feasible for HP^* . Moreover,*

$$y_{e(t)}^{(i)} b = c^* x_{e(t)}^{(i)} \xrightarrow[t \rightarrow T]{} \text{val}$$

(in fact, increases to val strictly monotonically) and the path $t \mapsto (y_{e(t)}^{(i)}, s_{e(t)}^{(i)})$ is bounded.

Proof. Assume $e \in \text{Swath}(i) \setminus \text{Core}(i)$. By Corollary 7.4, $x_e^{(i)}$ is optimal for the convex optimization problem

$$\begin{array}{ll} \min_x & -\ln c^*(e - x) - \frac{p_e^{(i)}(x)}{p_e^{(i+1)}(x)} \\ \text{s.t.} & Ax = b, \end{array}$$

and hence there exists $y = y_e^{(i)}$ satisfying the (rearranged) first-order condition

$$A^* y + s_e^{(i)} = c,$$

where, letting $q_e := p_e^{(i)} / p_e^{(i+1)}$ and $x_e = x_e^{(i)}$,

$$s_e^{(i)} := (c^*(e - x_e)) Dq_e(x_e) = \frac{c^*(e - x_e)}{p_e^{(i+1)}(x_e)} Dp_e^{(i)}(x_e),$$

(using $p_e^{(i)}(x_e) = 0$). Thus, to show the pair $(y_e^{(i)}, s_e^{(i)})$ is strictly feasible for HP^* , it suffices to show $s_e^{(i)} \in \text{int}(\Lambda_+^*)$, that is, assuming $z \in \Lambda_+$, $z \neq 0$, it suffices to show $Dq_e(x_e)[z] > 0$.

Corollary 4.2 shows $(\partial\Lambda_{+,e}^{(i)}) \setminus \Lambda_+$ contains no line segments of positive length other than those lying in rays $\{sx : s > 0\}$. Thus, the line segment connecting x_e to z lies entirely within $\Lambda_{+,e}^{(i)}$ with the exception of the point x_e and possibly the point z . Hence, $z(s) := (1 - s)z + sx_e$ satisfies $q_e(z(s)) > 0$ for $0 < s < 1$. Fix s strictly between 0 and 1.

Concavity of q_e (according to Theorem 7.2) and $q_e(x_e) = 0$ imply $q_e(z(s)) \leq Dq_e(x_e)[z(s) - x_e]$, whereas $q_e(x_e) = 0$ and homogeneity of q_e give $Dq_e(x_e)[x_e] = 0$. Thus,

$$0 < q_e(z(s)) \leq Dq_e(x_e)[z(s) - x_e] = (1 - s)Dq_e(x_e)[z],$$

completing the proof that $(y_e^{(i)}, s_e^{(i)})$ is strictly feasible for HP^* . Additionally observing

$$\begin{aligned} y_e^{(i)} b &= y_e^{(i)} A x_e \\ &= c^* x_e - s_e^{(i)} x_e \\ &= c^* x_e - (c^*(e - x_e)) Dq_e(x_e)[x_e] \\ &= c^* x_e, \end{aligned}$$

we thus see for a maximal trajectory $\{e(t) : 0 \leq t < T\}$, each pair $(y_{e(t)}^{(i)}, s_{e(t)}^{(i)})$ is strictly feasible for HP^* , and

$$\begin{aligned} y_{e(t)}^{(i)} b &= c^* x_{e(t)}^{(i)} \\ &\rightarrow \text{val} \quad \text{strictly monotonically} \quad (\text{by Proposition 8.2}). \end{aligned}$$

It only remains to show the path $t \mapsto (y_{e(t)}^{(i)}, s_{e(t)}^{(i)})$ is bounded, for which it suffices to show $t \mapsto s_{e(t)}^{(i)}$ is bounded (as A is surjective, by assumption). In turn, because $e(0) \in \Lambda_{++}$ and $s_{e(t)}^{(i)} \in \Lambda_+^*$, it suffices to show the value $s_{e(t)}^{(i)} e(0)$ is bounded from above independent of t . However,

$$\begin{aligned} s_{e(t)}^{(i)} e(0) &= c^* e(0) - y_{e(t)}^{(i)} A e(0) \\ &= c^* e(0) - y_{e(t)}^{(i)} b \\ &\rightarrow c^* e(0) - \text{val}, \end{aligned}$$

concluding the proof. \square

10. Proof of Theorem 6

We have now finished our analysis of the trajectories $t \mapsto e(t)$ (and paths $t \mapsto x_{e(t)}^{(i)}$) arising from the differential equation $\dot{e}(t) = x_{e(t)}^{(i)} - e(t)$ on the set $\text{Swath}(i) \setminus \text{Core}(i)$. However, in order to have a relatively complete picture of all of $\text{Swath}(i)$, results regarding the structure of $\text{Core}(i)$ are needed. That is the purpose of this section.

Specifically, just as the dynamics $\dot{e}(t) = x_{e(t)}^{(i)} - e(t)$ on $\text{Swath}(i) \setminus \text{Core}(i)$ leads to optimality, it would be nice as a matter of formalism to know that if $e \in \text{Core}(i)$ and $x \in \text{Opt}$, then moving e towards x results in a point $e + t(x - e)$ ($0 < t < 1$) also in $\text{Core}(i)$, thus retaining optimality.

Theorem 10.1. *Assume $1 \leq i \leq n - 2$. If $e \in \text{Core}(i)$ and $x \in \text{Opt}$ then*

$$\{e + t(x - e) \in \Lambda_{++} : t \in \mathbb{R}\} \subseteq \text{Core}(i).$$

Extending the theorem to include $i = 0$ is trivial, because $\text{Core}(0)$ is precisely the interior of Feas . It also is easy to extend the theorem to $i = n - 1$, simply because $\text{Core}(n - 1) = \emptyset$.⁷ We state the theorem only for $1 \leq i \leq n - 2$ to avoid having to be mindful of the special cases $i = 0, n - 1$ during the course of the proof.

By bootstrapping the theorem, the following elementary corollary provides some additional insight into the structure of $\text{Core}(i)$. The corollary is a restatement of Theorem 6.

⁷To see $\text{Core}(n - 1) = \emptyset$, recall that $(\partial\Lambda_{+,e}^{(n-1)}) \cap \Lambda_+$ is precisely the lineality space of Λ_+ (by Theorem 4.1(C)). Thus, since Λ_+ is regular (by assumption) and since the origin is infeasible (because $b \neq 0$), the boundary of Feas cannot intersect the boundary of $\text{Feas}_e^{(n-1)}$; in particular, it cannot happen that Opt intersects $\text{Opt}_e^{(n-1)}$.

Corollary 10.2. *If $e \in \text{Core}(i)$ and \mathcal{A} is the smallest affine space containing both e and Opt , then*

$$\mathcal{A} \cap \Lambda_{++} \subseteq \text{Core}(i) .$$

Proof. If Opt contains only a single point, then the corollary is nothing more than a restatement of the theorem. Thus, assume Opt contains more than a single point, and fix $\tilde{x} \in \text{relint}(\text{Opt})$. Then

$$\mathcal{A} = \{ \tilde{x} + \alpha(e - \tilde{x}) + \beta(x - \tilde{x}) : \alpha \in \mathbb{R}, \beta \geq 0 \text{ and } x \in \text{Opt} \} .$$

Assume $y \in \mathcal{A} \cap \Lambda_{++}$, and fix $\alpha \in \mathbb{R}, \beta \geq 0$ and $x \in \text{Opt}$ satisfying

$$y = \tilde{x} + \alpha(e - \tilde{x}) + \beta(x - \tilde{x}) .$$

To complete the proof, we show $y \in \text{Core}(i)$.

Since y and e lie in the interior of Feas, we have $c^*y, c^*e > \text{val} (= c^*\tilde{x}, c^*x)$, from which follows $\alpha > 0$. Thus, the point

$$z := \frac{\alpha}{\alpha+\beta}e + \frac{\beta}{\alpha+\beta}x$$

is well-defined, is a convex combination of e and x , and lies in Λ_{++} (because $e \in \Lambda_{++}$, $x \in \Lambda_+$ and $\alpha > 0$). As $e \in \text{Core}(i)$ and $x \in \text{Opt}$, Theorem 10.1 implies $z \in \text{Core}(i)$.

However, $y = z + t(\tilde{x} - z)$ for $t = \alpha + \beta - 1$. Hence, applying Theorem 10.1 again, now with z (resp., \tilde{x}) in place of e (resp., x), we find $y \in \text{Core}(i)$. \square

Before turning to the proof of Theorem 10.1, we recall the conjecture made in §2.

Conjecture. $\text{Core}(i)$ is convex.

It is highly unclear whether resolving the conjecture positively could have any relevance to algorithms. Nonetheless, a positive resolution would be interesting from a structural perspective.

Now we begin the process of proving Theorem 10.1. For motivation, we first consider a reasonably-general case for which the proof is rather straightforward, and afterwards move to the proof for the truly general case. That the proof is straightforward in one setting but not in full generality is reminiscent of the proof of Part I of the Main Theorem, where the case of “ $T = \infty$ ” was vastly easier than the case “ $T < \infty$ ” (i.e., §8.1 was a breeze compared to §8.2).

The straightforward case relies on the following lemma.

Lemma 10.3. *If $0 \leq i \leq n - 2$ and $x \in \partial\Lambda_{+,e}^{(i)}$, then*

$$Dp_e^{(i)}(x) = 0 \quad \Leftrightarrow \quad x \in \partial\Lambda_{+,e}^{(i+1)} .$$

Proof. This is Lemma 7 in [13]. \square

Here is the straightforward case:

Theorem 10.4. *Assume $1 \leq i \leq n - 2$ and $e \in \text{Core}(i)$. If $x \in \text{Opt} \cap \Lambda_{+,e}^{(i+1)}$ then*

$$\{e + t(x - e) \in \Lambda_{++} : t \in \mathbb{R}\} \subseteq \text{Core}(i) .$$

Proof. Recall Corollary 7.4, which shows that for every $\bar{e} \in \Lambda_{++}$ and $0 \leq i \leq n - 2$, the points lying in $\text{Opt}_{\bar{e}}^{(i)} \setminus \partial\Lambda_{+, \bar{e}}^{(i+1)}$ are precisely the optimal solutions to the linearly constrained, convex optimization problem

$$\begin{aligned} \min_{x \in \Lambda_{+, \bar{e}}^{(i+1)}} \quad & -\ln c^*(\bar{e} - x) - \frac{p_{\bar{e}}^{(i)}(x)}{p_{\bar{e}}^{(i+1)}(x)} \\ \text{s.t.} \quad & Ax = b , \end{aligned}$$

that is, for every $\bar{e} \in \Lambda_{++}$, letting $q_{\bar{e}}^{(i)} := p_{\bar{e}}^{(i)} / p_{\bar{e}}^{(i+1)}$,

$$\left. \begin{aligned} x &\in \text{Opt}_{\bar{e}}^{(i)} \cap \Lambda_{++,\bar{e}}^{(i+1)} \\ &\Leftrightarrow \\ \frac{1}{c^*(\bar{e}-x)} c^* - Dq_{\bar{e}}^{(i)}(x) &= y^* A \quad \text{for some } y^* \\ Ax &= b \\ x &\in \Lambda_{++,\bar{e}}^{(i+1)}. \end{aligned} \right\} \quad (10.1)$$

Now assume $x \in \text{Opt} \cap \Lambda_{++,\bar{e}}^{(i+1)}$, where $e \in \text{Core}(i)$ and $0 \leq i \leq n-2$. Then, since $\Lambda_+ \cap \partial\Lambda_{++,\bar{e}}^{(j)}$ ($0 \leq j \leq n-1$) is independent of $\bar{e} \in \Lambda_{++}$ (by Theorem 4.1(B)), we have for all $\bar{e} \in \Lambda_{++}$ that

$$x \in \Lambda_{++,\bar{e}}^{(i+1)}, \quad (10.2)$$

$$p_{\bar{e}}^{(j)}(x) = 0 \quad \text{for all } 0 \leq j \leq i, \quad (10.3)$$

and, using Lemma 10.3,

$$Dp_{\bar{e}}^{(j)}(x) = 0 \quad \text{for all } 0 \leq j \leq i-1. \quad (10.4)$$

Moreover, by (10.1), there exists y^* satisfying

$$\frac{1}{c^*(e-x)} c^* - Dq_e^{(i)}(x) = y^* A,$$

that is, as $p_e^{(i)}(x) = 0$, there exists y^* satisfying

$$\frac{1}{c^*(e-x)} c^* - \frac{1}{p_e^{(i+1)}(x)} Dp_e^{(i)}(x) = y^* A. \quad (10.5)$$

Fix t for which $e + t(x-e) \in \Lambda_{++}$ (thus, $-\infty < t < 1$). To prove the theorem, it suffices to show $x \in \text{Opt}_{e+t(x-e)}^{(i)}$, because then, $\text{Opt} \cap \text{Opt}_{e+t(x-e)}^{(i)} \neq \emptyset$, and hence, $\text{Opt} = \text{Opt}_{e+t(x-e)}^{(i)}$ (by Corollary 4.2(B)). Thus, in light of (10.2), we see from (10.1) that to prove the theorem, it suffices to show

$$\frac{1}{(1-t)c^*(e-x)} c^* - Dq_{e+t(x-e)}^{(i)}(x) = w^* A \quad \text{for some } w^*,$$

that is, as $p_{e+t(x-e)}^{(i)}(x) = 0$ (by (10.3)), it suffices to show

$$\frac{1}{(1-t)c^*(e-x)} c^* - \frac{1}{p_{e+t(x-e)}^{(i+1)}(x)} Dp_{e+t(x-e)}^{(i)}(x) = w^* A \quad \text{for some } w^*. \quad (10.6)$$

However,

$$\begin{aligned} Dp_{e+t(x-e)}^{(i)}(x) &= D^{i+1}p(x) \underbrace{[(1-t)e + tx, \dots, (1-t)e + tx]}_{i \text{ times}} \\ &= \sum_{j=0}^i \binom{i}{j} (1-t)^j t^{i-j} D^{i+1}p(x) \underbrace{[e, \dots, e]}_{j \text{ times}} \underbrace{[x, \dots, x]}_{i-j \text{ times}} \\ &= \sum_{j=0}^i \binom{i}{j} (1-t)^j t^{i-j} D^{i+1-j} p_e^{(j)}(x) \underbrace{[x, \dots, x]}_{i-j \text{ times}} \\ &= \sum_{j=0}^i \binom{i}{j} (1-t)^j t^{i-j} \frac{(n-j-1)!}{(n-i-1)!} Dp_e^{(j)}(x) \quad (\text{by (3.6)}) \\ &= (1-t)^i Dp_e^{(i)}(x) \quad (\text{by (10.4)}), \end{aligned}$$

and, similarly,

$$\begin{aligned} p_{e+t(x-e)}^{(i+1)}(x) &= \sum_{j=0}^{i+1} \binom{i+1}{j} (1-t)^j t^{i+1-j} \frac{(n-j)!}{(n-i-1)!} p_e^{(j)}(x) \\ &= (1-t)^{i+1} p_e^{(i+1)}(x) \quad (\text{by (10.3)}) . \end{aligned}$$

Hence, $w^* = \frac{1}{1-t} y^*$ satisfies (10.6) (where y^* is as in (10.5)), completing the proof. \square

In the course of the proof for the “truly general” case, there is a crucial outside result to which we refer, and thus before formally beginning the proof, we explain the result.

For $e \in \Lambda_{++}$ and an arbitrary point x , let $M_e(x)$ denote the number of non-negative roots of $t \mapsto p(x+te)$. The result is that $M_e(x)$ is independent of $e \in \Lambda_{++}$, as was stated in [9] (as Theorem 2.12). (In [13], the fact was (essentially) established *only* for $x \in \Lambda_+$.)

Perhaps worth recording is that the independence easily follows from a most useful tool in the hyperbolic polynomial literature:

Helton-Vinnikov Theorem. *Assume $e \in \Lambda_{++}$. For any points x and z , there exist $n \times n$ symmetric matrices X and Z such that*

$$(r, s, t) \mapsto p(rx + sz + te) = p(e) \det(rX + sZ + tI) .$$

This formulation of the theorem comes from [11], where it was obtained by straightforward homogenization of the original result in [10]. (The initial importance of the homogeneous version was that it affirmatively settled the “Lax Conjecture” – see [11] for discussion.) (See [3] for negative results on possibilities of extensions to more than three variables r, s, t .)

Corollary 10.6. *For $e \in \Lambda_{++}$ and arbitrary x , let $M_e(x)$ denote the number of non-negative roots for the polynomial $t \mapsto p(x+te)$. The value $M_e(x)$ is the same for all $e \in \Lambda_{++}$.*

Proof. First observe that because all roots of univariate polynomials vary continuously in the coefficients so long as the leading coefficient does not vanish, it suffices to show $\text{mult}_e(x)$ is independent of $e \in \Lambda_{++}$, where $\text{mult}_e(x)$ is the multiplicity of 0 as a root of $t \mapsto p(x+te)$. However, for $e, z \in \Lambda_{++}$, using the Helton-Vinnikov Theorem,

$$\text{mult}_e(x) = n - \text{rank}(X) = n - \text{rank}(Z^{-1/2} X Z^{-1/2}) = \text{mult}_z(x) ,$$

and thus the corollary is established. \square

Proof of Theorem 10.1. Fix $e \in \text{Core}(i)$, where $1 \leq i \leq n-2$. Let x denote a point in Opt ($= \text{Opt}_e^{(i)}$).

Let $e(t) := e + t(x-e)$, and let I denote the open interval consisting of all $t \in \mathbb{R}$ for which $e(t) \in \Lambda_{++}$ – thus, $I \subseteq (-\infty, 1)$.

Since $\Lambda_+ \cap \partial\Lambda_{+, \bar{e}}^{(i)}$ is independent of $\bar{e} \in \Lambda_{++}$ (by Corollary 4.2(A)), and since $\text{Opt}_e^{(i)} = \text{Opt}$, we have

$$\text{Opt} \subset \partial\Lambda_{+, e(t)}^{(i)} \text{ for all } t \in I . \quad (10.7)$$

Let $\mathcal{D} := \{d : Ad = 0 \text{ and } c^*d = 0\}$, and let $x + \mathcal{D}$ be the set consisting of all points $x + d$ where $d \in \mathcal{D}$.

Since $\text{relint}(\text{Feas}_{e(t)}^{(i)}) = (x + \{d : Ad = 0\}) \cap \Lambda_{++, e(t)}^{(i)}$ and $x \in \text{Opt} \subset \text{Feas}_{e(t)}^{(i)}$, we have

$$x \in \text{Opt}_{e(t)}^{(i)} \quad \Leftrightarrow \quad (x + \mathcal{D}) \cap \Lambda_{++, e(t)}^{(i)} = \emptyset .$$

On the other hand, x is in $\text{Opt}_{e(t)}^{(i)}$, as well as in Opt , if and only if $e(t) \in \text{Core}(i)$ (by Corollary 4.2(B)). Consequently,

$$e(t) \in \text{Core}(i) \quad \Leftrightarrow \quad (x + \mathcal{D}) \cap \Lambda_{++,e(t)}^{(i)} = \emptyset.$$

Thus, our goal – proving $e(t) \in \text{Core}(i)$ for all $t \in I$ – will be accomplished if we prove $(x + \mathcal{D}) \cap \Lambda_{++,e(t)}^{(i)} = \emptyset$ for all $t \in I$. Hence, fixing arbitrary $d \in \mathcal{D}$, letting $\text{ray}(d) := \{sd : s \geq 0\}$ and

$$I(d) := \{t \in I : (x + \text{ray}(d)) \cap \Lambda_{++,e(t)}^{(i)} = \emptyset\},$$

our goal is to show $I(d) = I$.

Recall, however, the characterization (3.2) now applied to $\Lambda_{++,e(t)}^{(i)}$:

$$\Lambda_{++,e(t)}^{(i)} := \{x : p_{e(t)}^{(j)}(x) > 0 \text{ for all } j = i, \dots, n-1\}. \quad (10.8)$$

Clearly, then, $I \setminus I(d)$ is an open subset of the open interval I . Hence, since $I(d) \neq \emptyset$ (indeed, $0 \in I(d)$), to complete the proof of the theorem, it suffices to show that $I(d)$ is open. As e is only assumed to be an element of $\text{Core}(i)$ – in particular, e could be replaced with any $e(t)$ that happens to be in $\text{Core}(i)$ – our goal has become:

$$\text{Show } I(d) \text{ contains an open interval including } 0. \quad (10.9)$$

For non-negative integers j and k , define

$$\alpha_{jk}(t) := D^k p_{e(t)}^{(j)}(x) \underbrace{[d, \dots, d]}_{k \text{ times}}$$

and

$$\beta_j(t) := \begin{cases} 0 & \text{if } \alpha_{jk}(t) = 0 \text{ for all } k; \\ \alpha_{j,k(j,t)} & \text{otherwise, where } k(j,t) := \min\{k : \alpha_{jk}(t) \neq 0\}. \end{cases}$$

Note that if $\beta_j(t) = 0$, then the polynomial $s \mapsto p_{e(t)}^{(j)}(x + sd)$ is identically zero, whereas if $\beta_j(t) \neq 0$, the polynomial evaluated at small, positive s has the same sign as $\beta_j(t)$. Thus, since x lies in the convex set $\Lambda_{+,e(t)}^{(i)}$, we have by the characterization (10.8) that

$$(x + \text{ray}(d)) \cap \Lambda_{++,e(t)}^{(i)} \neq \emptyset \quad \Leftrightarrow \quad \beta_j(t) > 0 \text{ for all } j = i, \dots, n-1.$$

Our goal (10.9) has now been reduced to:

$$\text{Show } \exists \epsilon > 0 \text{ such that } |t| < \epsilon \Rightarrow \beta_j(t) \leq 0 \text{ for some } j \in \{i, \dots, n-1\}. \quad (10.10)$$

For this we consider two cases.

First, assume $\beta_j(0) \geq 0$ for all $j = i, \dots, n-1$. Then $x + sd \in \text{Feas}_e^{(i)}$ for all sufficiently small, positive s . Thus, since $x \in \text{Opt}_e^{(i)} = \text{Opt}$ and $c^*d = 0$, we have $x + sd \in \text{Opt}$ for all sufficiently small, positive s . Hence, by (10.7), for each $t \in I$, the univariate polynomial $s \mapsto p_{e(t)}^{(i)}(x + se)$ has value zero on an open interval, and thus is identically zero, so $\beta_i(t) = 0$ for all $t \in I$, (more than) accomplishing (10.10) for the first case.

Now consider the remaining case, that is, assume $\beta_j(0) < 0$ for some $j \in \{i, \dots, n-1\}$; fix such a j . To accomplish (10.10), of course it suffices to show for this fixed value of j that $\beta_j(t) < 0$ for all t in an open interval containing 0. In turn, by definition of $\beta_j(t)$, it suffices to:

$$\text{Show } \exists \epsilon > 0 \text{ such that } |t| < \epsilon \Rightarrow D^{k(j,0)} p_{e(t)}^{(j)}(x) \underbrace{[d, \dots, d]}_{k(j,0) \text{ times}} < 0 \quad (10.11)$$

and

$$\text{show } (t \in I) \wedge (k < k(j, 0)) \Rightarrow D^k p_{e(t)}^{(j)}(x) \underbrace{[d, \dots, d]}_{k \text{ times}} = 0. \quad (10.12)$$

The existence of ϵ as in (10.11) is simply a matter of continuity and $\beta_j(0) < 0$. For accomplishing (10.12), observe

$$\begin{aligned} D^k p_{e(t)}^{(j)}(x) &= D^{k+j} p(x) \underbrace{[(1-t)e + tx, \dots, (1-t)e + tx]}_{j \text{ times}} \\ &= \sum_{\ell=0}^j \binom{j}{\ell} (1-t)^\ell t^{j-\ell} D^{k+j-\ell} p_e^{(\ell)}(x) \underbrace{[x, \dots, x]}_{\ell-j \text{ times}} \\ &= \sum_{\ell=0}^j \binom{j}{\ell} (1-t)^\ell t^{j-\ell} \frac{(n-k-\ell)!}{(n-k-j)!} D^k p_e^{(\ell)}(x) \quad (\text{using (3.6)}). \end{aligned}$$

Consequently, (10.12) is immediately accomplished by the following proposition, thus concluding the proof of the theorem (except for proving the proposition). \square

Proposition 10.7. *Assume $e \in \Lambda_{++}$, $x \in \Lambda_+$ and let d be a vector. Assume non-negative integers j and k satisfy $D^k p_e^{(j)}(x) \underbrace{[d, \dots, d]}_{k \text{ times}} \neq 0$; let $k(j)$ be the smallest*

such k for j and assume $k(j) > 0$.

Then

$$D^k p_e^{(\ell)}(x) \underbrace{[d, \dots, d]}_{k \text{ times}} = 0 \quad \text{for all } \ell = 0, \dots, j \text{ and } k = 0, \dots, k(j) - 1.$$

The proof of the proposition makes use of the following lemma.

For $e \in \Lambda_{++}$ and arbitrary z , recall that $M_e(z)$ denotes the number of non-negative roots of $t \mapsto p(z + te)$.

Lemma 10.8. *Assume $e, x \in \Lambda_{++}$, let $d \neq 0$, and let $s_1 \leq s_2 \leq \dots \leq s_k$ denote the positive roots (including multiplicities) of $s \mapsto p(x + sd)$. Then, for every $\bar{s} \geq 0$,*

$$M_e(x + \bar{s}d) = \#\{i : s_i \leq \bar{s}\}.$$

Proof. Because $M_e(x + \bar{s}d)$ is independent of $e \in \Lambda_{++}$ (by Corollary 10.6), in proving the lemma we may assume $e = x$, in which case we need only consider the hyperbolic polynomial obtained by restricting p to the subspace spanned by x and d . However, every hyperbolic polynomial in two variables of degree n is of the form $(y_1, y_2) \mapsto \prod_{j=1}^n a_j^T y$ for some vectors $a_j \in \mathbb{R}^2$ (this easily follows from two facts: (i) every complex homogeneous polynomial in two variables of degree n is of the form $(y_1, y_2) \mapsto \prod_{j=1}^n (a_{j,1}y_1 + a_{j,2}y_2)$ for some $a_j \in \mathbb{C}^2$; (ii) if $a \in \mathbb{C}^2$ is not a (complex) multiple of a real vector, then $\{(y_1, y_2) \in \mathbb{R}^2 : a_1y_1 + a_2y_2 = 0\}$ contains only the origin).

Thus, we need only consider hyperbolic polynomials $p(y_1, y_2) = \prod_{j=1}^n a_j^T y$ – where $a_j \in \mathbb{R}^2$ – and $x = (x_1, x_2)$ which for all j satisfies $a_j^T x \neq 0$. For $(d_1, d_2) \neq (0, 0)$, our goal is to show for $\bar{s} \geq 0$ that the number of roots $0 < \hat{s} \leq \bar{s}$ (counting multiplicities) for the univariate polynomial

$$s \mapsto \prod_{j=1}^n a_j^T (x + sd) \quad (10.13)$$

is the same as the number of roots $\hat{t} \geq 0$ (counting multiplicities) for

$$t \mapsto \prod_{j=1}^n a_j^T (x + \bar{s}d + tx). \quad (10.14)$$

Of course, however, $\hat{s} > 0$ is a root of (10.13) if and only if $\hat{s} = \bar{s}/(1 + \hat{t})$ for some root $\hat{t} \geq 0$ of (10.14). The lemma follows. \square

Proof of Proposition 10.7 Assume $e \in \Lambda_{++}$. For arbitrary z and $0 \leq \ell \leq n-1$, let $M_e^{(\ell)}(z)$ denote the number of non-negative roots (counting multiplicities) for $t \mapsto p_e^{(\ell)}(x + te)$. Additionally, for a vector d and value $s \geq 0$, let $N_e^{(\ell)}(x, d, s)$ denote the number of roots (counting multiplicities) in the closed interval $[-s, s]$ for the univariate polynomial $\bar{s} \mapsto p_e^{(\ell)}(x + \bar{s}d)$; if the univariate polynomial is identically zero, let $N_e^{(\ell)}(x, d, s) = \infty$.

If $x \in \Lambda_{++}$ (hence $x \in \Lambda_{++}^{(\ell)}$ for all $0 \leq \ell \leq n-1$), $d \neq 0$ and $s \geq 0$, Lemma 10.8 can be applied with $p_e^{(\ell)}$ in place of p , and applied with $-d$ as well as with d , yielding

$$N_e^{(\ell)}(x, d, s) = M_e^{(\ell)}(x + sd) + M_e^{(\ell)}(x - sd). \quad (10.15)$$

On the other hand, for any z , the interlacing of the roots of $t \mapsto p_e^{(\ell)}(z + te)$ and its derivative $t \mapsto p_e^{(\ell+1)}(z + te)$ gives $M_e^{(\ell)}(z) \geq M_e^{(\ell+1)}(z)$, and thus,

$$\ell \leq j \quad \Rightarrow \quad M_e^{(\ell)}(z) \geq M_e^{(j)}(z). \quad (10.16)$$

From (10.15) and (10.16) follows

$$(x \in \Lambda_{++}) \wedge (\ell \leq j) \quad \Rightarrow \quad N_e^{(\ell)}(x, d, s) \geq N_e^{(j)}(x, d, s). \quad (10.17)$$

Assume, now, that x, d, j and $k(j)$ satisfy the hypothesis of the proposition. Observe that definitions readily give $k(j) = N_e^{(j)}(x, d, 0)$. Moreover, proving the proposition amounts precisely to showing

$$N_e^{(\ell)}(x, d, 0) \geq N_e^{(j)}(x, d, 0) \quad \text{for } \ell = 0, \dots, j. \quad (10.18)$$

For $\epsilon > 0$, let $x(\epsilon) := x + \epsilon e$, a point in Λ_{++} . Each polynomial $\bar{s} \mapsto p_e^{(\ell)}(x(\epsilon) + \bar{s}d)$ ($0 \leq \ell \leq n-1$) is not identically zero and has only real roots (indeed, by homogeneity, the roots are the reciprocals of the non-zero roots for $t \mapsto p_e^{(\ell)}(d + tx(\epsilon))$). Thus, since

$$\left. \begin{array}{l} \text{bounded roots of a univariate polynomial} \\ \text{vary continuously in the coefficients} \\ \text{(so long as the polynomial does not become identically zero)} \end{array} \right\} \quad (10.19)$$

we have for $0 \leq \ell \leq j$ that either $N_e^{(\ell)}(x, d, 0) = \infty$ (i.e., $\bar{s} \mapsto p_e^{(\ell)}(x + \bar{s}d) \equiv 0$), or

$$\begin{aligned} N_e^{(\ell)}(x, d, 0) &= \lim_{s \downarrow 0} \left(\lim_{\epsilon \downarrow 0} \left(N_e^{(\ell)}(x(\epsilon), d, s) \right) \right) \\ &\geq \lim_{s \downarrow 0} \left(\lim_{\epsilon \downarrow 0} \left(N_e^{(j)}(x(\epsilon), d, s) \right) \right) \quad (\text{by (10.17)}) \\ &\geq \lim_{s \downarrow 0} \left(N_e^{(j)}(x, d, s/2) \right) \quad (\text{again using (10.19)}) \\ &= N_e^{(j)}(x, d, 0), \end{aligned}$$

thereby establishing (10.18) and hence completing the proof. \square

REFERENCES

- [1] H. H. Bauschke, O. Güler, A. S. Lewis and H. S. Sendov, Hyperbolic polynomials and convex analysis, *Canadian Journal of Mathematics* **53** (2001) no. 3, 470–488.
- [2] J. Borcea and P. Brändén, Multivariate Pólya-Schur classification problems in the Weyl algebra, *Proceedings of the London Mathematical Society* **101** (2010) no. 1, 73–104.
- [3] P. Brändén, Obstructions to determinantal representability, *Advances in Mathematics* **226** (2011) no. 2, 1202–1212.
- [4] C.B. Chua, The primal-dual second-order cone approximations algorithm for symmetric cone programming, *Foundations of Computational Mathematics* **7** (2007), no. 3, 271–302.

- [5] A.V. Fiacco and G.P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, SIAM, Philadelphia, PA, 1990.
- [6] L. Gårding, An inequality for hyperbolic polynomials, *Journal of Mathematics and Mechanics* **8** (1959) no. 6, 957–965.
- [7] O. Güler, Hyperbolic polynomials and interior point methods for convex programming, *Mathematics of Operations Research* **22** (1997), no. 2, 350–377.
- [8] L. Gurvits, Van der Waerden/Schrijver-Valiant like conjectures and stable (aka hyperbolic) homogeneous polynomials: one theorem for all, *Electronic Journal of Combinatorics* **15** (2008) #R66.
- [9] F.R. Harvey and H.B. Lawson Jr., Hyperbolic polynomials and the Dirichlet problem, preprint available at arXiv.org.
- [10] J. Helton and V. Vinnikov, Linear matrix inequality representations of sets, *Communications of Pure and Applied Mathematics* **60** (2007) 654–674.
- [11] A.S. Lewis, P.A. Parrilo, M.V. Ramana, The Lax conjecture is true, *Proceedings of the American Mathematical Society* **133** (2005) no. 9, 2495–2499.
- [12] Y. Nesterov and A. Nemirovski, *Interior-point polynomial algorithms in convex programming*, SIAM, Philadelphia, PA, 1994.
- [13] J. Renegar, Hyperbolic programs, and their derivative relaxations, *Foundations of Computational Mathematics* **6** (2006) no. 1, 59–79.
- [14] Y. Zinchenko, *The local behavior of the Shrink-Wrapping algorithm for linear programming*, Ph.D. thesis, Cornell University, 2005.
- [15] Y. Zinchenko, On hyperbolicity cones associated with elementary symmetric polynomials, *Optimization Letters* **2** (2008) no. 3, 389–402.
- [16] Y. Zinchenko, Shrink-wrapping trajectories for linear programming, preprint available at optimization-online.org.

SCHOOL OF OPERATIONS RESEARCH AND INFORMATION ENGINEERING, CORNELL UNIVERSITY, ITHACA, NY, U.S.